Measurement Of Engagement Rate On Instagram For Business Marketing (Case Study: MSME of Dowry In Jember)

Putranto, Hermawan Arief[1]
Departement of Information
Technology
State Polytehnic of Jember
Jember, Indonesia
hermawan_arief@polije.ac.id

Rizaldi, Taufiq[2]
Departement of Information
Technology
State Polytehnic of Jember
Jember, Indonesia
taufiq_r@polije.ac.id

Riskiawan, Hendra Yufit[3]
Departement of Information
Technology
State Polytehnic of Jember
Jember, Indonesia
yufit@polije.ac.id

Setyohadi, Dwi Putro Sarwo[4]
Departement of Information
Technology
State Polytehnic of Jember
Jember, Indonesia
dwi.putro@polije.ac.id

Atmadji, Ery Setiyawan Jullev[5]
Departement of Information
Technology
State Polytehnic of Jember
Jember, Indonesia
ery@polije.ac.id

Nuryanto, Ishaq Husainy[6]
Departement of Information
Technology
State Polytehnic of Jember
Jember, Indonesia

Abstract— **Based on data released by APJII in 2020, internet users in Indonesia which reached 73.7% of the Indonesian population are directly proportional to marketing activities through the internet, one of which is through the social media platform Instagram. Many business people communicate with their prospective customers through Instagram, one of which is the delivery MSMEs in Jember Regency. One of the benchmarks for measuring the level of interaction between MSMEs and potential consumers is through the engagement rate. We carry out a data mining process to track and summarize engagement activities on delivery MSME Instagram accounts in Jember district by utilizing a scraping technique based on the PHP programming language which is connected to the RESTful Application Programming Interface (API) provided by Instagram. From the interpolation calculation based on the correlation between the number of followers and the average value of the engagement rate, we categorize the standard value of the average engagement rate according to the case study of this research. From the evaluation results, it was found that the majority of accounts still have an average engagement value below the ideal value with a deviation value around -6.39%.**

Keywords—Data Mining, Engagement Rate, Scraping, Social Media

I. INTRODUCTION

The rapid development of information technology has an impact on increasing internet users in Indonesia. APJII noted that in 2020 internet users in Indonesia reached 196.71 million people or 73.7% of the total population of Indonesia [1]. Most internet users are used to access existing content on social media where the most frequently accessed social media in a row are Facebook, Instagram, linkedin and twitter [1]. Apart from increasing the number of active internet users, the covid-19 pandemic also increases the number of buying and selling activities through online media or social media, like Instagram. Since 2019, there are around 62 million active users in Indonesia, therefor, the platform is widely used to market products and services, especially by MSMEs, where marketing through Instagram can increase consumer buying interest up to 50.2% [2]. The potential for marketing through social media, especially in Instagram, is unfortunately not directly proportional to the understanding of MSME actors so that it is not optimal in using these features to support marketing activities.

An understanding of the Engagement Rate (ER) in a marketing activity will help business actors to map marketing potential using social media. The utilization of the Instagram API to determine the ER value on the accounts of Indonesian government ministries and agencies shows that 50.82% of them were not able to optimize their Instagram accounts to achieve engagement from their followers [3][4]. A special algorithm can be applied to determine the efficiency of influencers for a product they introduce. This can be used to optimize the selection of influencers to increase the Engagement Rate which has a positive impact on the marketing of a product [5] [6]. In addition, the Engagement Rate can also be used to determine which articles or photos will be published, so that information distribution becomes more effective [4] [7].

In this study, data was collected using the Web Scrapping method by Instagram API and simple html DOM library, then the data was processed using the Instagram Engagement matrix formula [5] which will produce an analysis related to the relationship between the number of followers, the number of comments and likes per post that will be measured so that it can calculate the average value of the engagement rate. This calculation further analyzed the relationship between the number of followers, comments and likes to the value of the engagement rate. In the next chapter, related work will be discussed in terms of calculating the value of the engagement rate. In chapter three, we will discuss the data collection process that will be used to calculate the value of the engagement rate metric.

II. RELATED WORK

The use of ER to see the level of community participation in an idea or opinion has been reviewed in several studies. Among them is a study conducted by Arman et al [3], who proposed a new formula to measure the level of citizen engagement on the Instagram of Indonesian government ministries and agencies. From the research, they found that many countries have used social media to communicate with their people, and even though Indonesia has the largest population, Indonesia's E-participation Index (EPI) is still

ranked 92 with a value of 0.618. This value is far below Malaysia, the Philippines and Singapore. The approach taken is the use of data mining to track and summarize engagement activities on the Instagram accounts of Ministries and Institutions for one year, between 1 January - 31 December 2018. The results obtained from the comparative evaluation show that the new formula proposed uses more variables related to Engagement Rate. (ER) so that it is more representative in describing the ER variable. In addition, in this study it was also found that the higher the number of followers, the lower the ER value, so that the grouping is in the form of a range of ER values based on the number of followers. So, it can be concluded that the high number of followers cannot indicate the level of participation of the account owner and his followers.

The second study also views that social media has become an important part of governance of government institutions in the Digital Era 4.0 [8], one of which is in Health Institutions. By using Instagram, which is one of the social media used by Health Institutions, using the level of reader involvement as a parameter in measuring the effectiveness of each health office in disseminating health information. This study aims to analyze the level of reader participation by measuring the ER scores of Instagram accounts of several government health agencies. This study uses web scraping techniques to collect data from Instagram such as the number of followers, likes, comments, and posts. The results of this study, can find out which government institutions can reach their followers the most, judging from their ER level.

Another study conducted by Liao et al [9] also used ER to maximize the value of the marketing budget. This is one of the priority issues for the company. The proposed approach is to effectively analyze Facebook photo posts, so that companies and social media managers can concentrate on handling their fan pages. This study aims to use text mining techniques to find audiences accurately. They built a topic model recommendation system (TMRS) to analyze Facebook posts by sorting target posts according to their recommended scores. Based on the results of the recommendations, it was verified that TMRS can increase the level of engagement compared to the traditional recommended level of engagement (ERRM) method. The result, from the proposed approach shows that the company has the opportunity to create advertising with a limited budget but has a high participation rate.

From several studies that have been carried out, it is certain that ER can be used to see the level of participation between followers and the social media accounts they follow. This level of participation is also positively correlated with the reach of social media accounts, so it is appropriate if ER is also used to analyze the extent to which a product can be accepted by the public.

III. RESULT AND DISCUSSION

A. Pengambilan Data

This study uses the Instagram Public (Authentication) API to access posts of an account, by utilizing the Instagram Scraper endpoints, the required data, such as images, videos, captions, hashtags, comments, and likes will be obtained from a target account. The data retrieval process uses the Document Object Model (DOM) and Client URL (cURL). By using it, the application will have a way to access the content, structure and style of HTML documents using JavaScript in simple HTML DOM library. cURL is used for the scraping process, the data obtained is entered into a new variable and then extracted using PHP Simple HTML DOM.
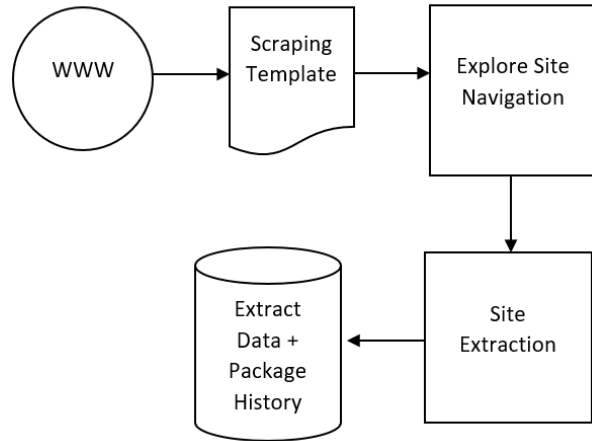

Fig. 1. Data Collection Flow

These functions are packaged in an interface to simplify the scarping process. To carry out the analysis process, it is necessary to enter a non-private Instagram account for which the ER matrix will be calculated.
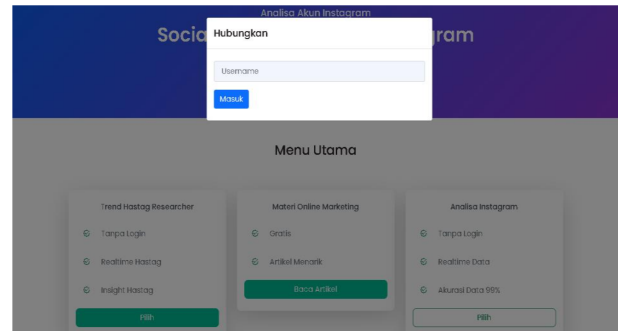

Fig. 2. Instagram account analysis feature interface

Instagram provides a RESTful Application Programming Interface (API) which can be accessed by making a request to the Instagram server. By utilizing the Instagram API, the latest posts from the account will be taken along with the number of likes, comments and the results of calculating the engagement rate for each post, both in the form of images and videos from the account. The amount of data retrieved is determined based on the value inserted in the program code function chart. There are 10 responses obtained from the system such as status, id, shortcode, has_like, has_commented, thumbnail, video, url, media_type, engagement. The request data obtained is converted from JSON form into a CSV file for each account using the PHP programming language. The data in the form of JSON is also displayed on a web-based display as shown in Figure 3.
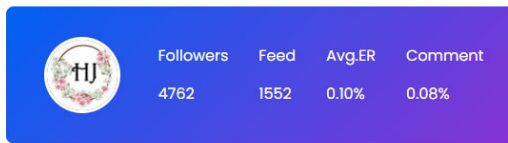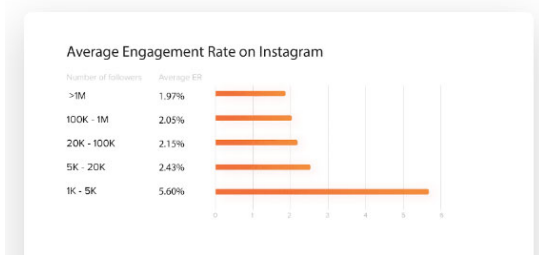
Fig. 3. Account Post Collection Results and ER Calculations

B. Matrix Engagement Rate

In this research, there are two formulas which is used based on the literature study shown in table 1.

TABLE I. FORMULA BERDASARKAN STUDI LITERATUR

No	Formula	Literature Study
1	Instagram Engagement Rate	Ryhänen [10]
2	Average Engagement Rate	Primasiwi, Irawan, & Ambarwati [5]

The Engagement Rate Matrix is a calculation matrix used on social media accounts. ER matrix will calculate an account's interactions and impressions based on the number of followers for each content or post created by an account. The results of this Engagement Rate calculation will be displayed in the form of a percentage obtained from the interaction of an account with its followers.

For the calculation of the Instagram Engagement Rate, use following formula:

$$Instagram\ Er(\%) = \frac{L+C}{F} \times 100 \quad (1)$$

Where:

Er = Engagement Rate

L = Likes

C = Comment

F = Followers

After we get the ER percentage results from each account post, then, we calculate the average Engagement Rate against x number of posts in that account. Which is x is a value that can be adjusted through the program code. To calculate the Average Instagram Engagement Rate, use the following formula:

$$Average\ Er(\%) = \frac{\frac{L+C}{F}}{Cn} \times 100$$

Where:

Er = Engagement Rate

L = Likes

C = Comment

F = Followers

Cn = Total Post standard average ER

After the average value of the Engagement Rate has been obtained, then this value is compared with the average value of the ideal Engagement Rate which can be seen in table 2. This table contains the ideal average value of ER which is generally obtained by an account based on the number of followers [3] [10]. In each follower number range, there is an ideal percentage for the Average Engagement Rate value, if the percentage value obtained is still below the expected, it can be said that the level of interaction between the account and its followers is still lacking. In this study, we formulate a different assessment than the one shown in table 2 [10].

TABLE II. IDEAL AVERAGE ENGAGEMENT RATE

No.	Number of Followers	Average ER
1.	> 1 M	1.54 %
2.	100 K – 1 M	1.55 %
3.	20 K – 100 K	1.62 %
4.	5 K – 20 K	2.28 %
5.	1 K – 5K	4.26 %

In some cases, it was found that the correlation analysis on the ER value showed that the number of followers was not positively correlated with a high ER value [3], based on this an interpolation calculation was carried out to determine a new range of engagement rate values where the number of followers was adjusted to the data that was obtained. The results of the interpolation calculation are proven in table III where the interpolation value is used to see the ER value that corresponds to the number of followers.

TABLE III. INTERPOLASI ANALISIS RESULT

Values	Forecast	Lower Confidence Bound	Upper Confidence Bound
1,72	-		
2,37	-		
2,65	-		
4,98	-		
7,58	7,58	7,58	7,58
-	8,5377259	6,64	10,43
-	10,043023	7,92	12,16
-	11,54832	9,23	13,87
-	13,053617	10,54	15,56
-	14,558915	11,87	17,24

Based on the results of the interpolation analysis, we propose a different assessment as shown in table IV where the average engagement rate value is grouped into 8 groups with the number of followers calculated from zero (0) to five hundred thousand (500.000).

TABLE IV.   NEW RANGE OF AVERAGE ENGAGEMENT RATE

	Follower	Avg ER
1	0 - 100	11,54
2	101 - 500	10,04
3	501 - 1000	8,53
4	1001 - 5000	7,58
5	5001 - 20000	4,98
6	20001 - 100000	2,65
7	100001 - 300000	2,37
8	300001 - 500000	1,72

C. Measurement of Engagement Rate

In a study to measure the Engagement Rate on MSMEs of dowry in Jember. First, we carried out the process of collecting account data from MSMEs who were members of Seni Mahar Jember, which is an organization that protects workers' associations or MSMEs engaged in the dowry sector. The second step is data scraping process from Instagram and calculating the average engagement activities where the average value will be measured based on the range of average ER values as shows in table IV. The calculation process begins by calculating the ER value for each post and then looking for the average value of all posts. The results of the data retrieval process are shown in table V.

TABLE V.   ENGAGEMENT RATE MEASUREMENT RESULTS

No	Instagram Account	Follower	Post	Feed	Av ER (%)
1	hadiahimpianku	66	10	57	8.19%
3	weddinggalleryjember2	69	10	720	0.00%
4	etflowers.id	956	10	337	0.44%
5	ayune_seserahan	58	10	21	15.17%
6	rubia_gallery	520	10	426	0.13%
7	kirey.seserahan	105	10	45	2.67%
8	hantaran_jember	277	10	248	0.11%
10	binarseserahan	98	10	31	4.90%
11	daffi.seserahan	279	10	55	7.06%
12	hantaranjember	4783	10	1552	0.11%
13	lunariagift	598	10	16	3.38%
14	seserahan_bybarra	1232	10	1066	0.31%
15	eimoi.id	531	10	164	0.85%
16	naviamates	1213	10	59	0.25%
17	maharjember	296	10	143	2.60%
18	mahar.jember	408	3	3	0.49%
19	hantaranku.gumush	9	10	12	12.22%
20	ini.maharku	60	10	85	0.67%
21	_maharbintang_	30	10	18	10.33%
22	hantarankaukah.jember	386	10	241	2.67%
23	se.ri.us.id	453	10	45	1.61%
24	defaco.id	2786	5	5	0%
25	newtomorrow.jember	674	10	45	0.68%
26	cantiknyamahar	11383	10	1777	0.41%
27	weddinggalleryjember	4176	10	8935	0.06%
28	griyaayudya	4118	10	512	1.14%
29	inewenihandmade	8181	10	3384	2.18%
30	lynds.creative	1750	10	2814	0.25%
31	ainirumahhandmade	2000	10	1901	0.17%

As shown in table V, the highest ER was obtained by the ayune_seserahan account with a value of 15.17% and the lowest ER value was obtained by the devaco.id account with an average ER of 0%. Meanwhile, the number of followers owned by each account is 58 and 2786 respectively. This is of course very different from the average ER value obtained. The same data is also shown by the beautiful dowry account which has the highest number of followers, 11383 only having an average ER value of 0.41%. This proves that the number of followers is not positively correlated with the average ER value obtained.

However, it does not mean that a high average ER value also means that it shows good interaction between account owners and their followers. Each ER mean value should be compared with the ideal ER value in table IV. Accounts that get good results for their ER quality mean that the average ER value of the account is in the standard ER value range and indicates that the account is able to optimize the account so that followers are willing to participate in every post, the opposite is an account that has a value below the average ER standard range indicates that the account has not been able to optimize the post to attract followers to participate.

TABLE VI.   ENGAGEMENT RATE MEASUREMENT RESULTS WITH IDEAL ENGAGEMENT RATE

No	Instagram Account	Av ER (%)	Ideal Av ER	Deviation
1	hadiahimpianku	8.19%	11.54%	-3,35%
3	weddinggalleryjember2	0.00%	11.54%	-11,54%
4	etflowers.id	0.44%	8.53%	-8,09%
5	ayune_seserahan	15.17%	11.54%	3,63%
6	rubia_gallery	0.13%	8.53%	-8,4%
7	kirey.seserahan	2.67%	10.04%	-7,37%
8	hantaran_jember	0.11%	10.04%	-9,93%
10	binarseserahan	4.90%	11.54%	-6,64%
11	daffi.seserahan	7.06%	10.04%	-2,98%
12	hantaranjember	0.11%	7.58%	-7,47%
13	lunariagift	3.38%	8.53%	-5,15%
14	seserahan_bybarra	0.31%	7.58%	-7,27%
15	eimoi.id	0.85%	8.53%	-7,68%
16	naviamates	0.25%	7.58%	-7,33%
17	maharjember	2.60%	10.04%	-7,44%
18	mahar.jember	0.49%	10.04%	-9,55%
19	hantaranku.gumush	12.22%	11.54%	0,68%
20	ini.maharku	0.67%	11.54%	-10,87%
21	_maharbintang_	10.33%	11.54%	-1,21%
22	hantarankaukah.jember	2.67%	10.04%	-7,37%
23	se.ri.us.id	1.61%	10.04%	-8,43%

24	defaco.id	0%	7.58%	-7,58%
25	newtomorrow.jember	0.68%	8.53%	-7,85%
26	cantiknyamahar	0.41%	4.98%	-4,57%
27	weddinggalleryjember	0.06%	7.58%	-7,52%
28	griyaayudya	1.14%	7.58%	-6,44%
29	inewenihandmade	2.18%	4.98%	-2,8%
30	lynds.creative	0.25%	7.58%	-7,33%
31	ainirumahhandmade	0.17%	7.58%	-7,41%

In table VI it can be seen that positive deviations only occur in two accounts, ayune_seserahan and hantaranku.gumush, regardless of the number of followers and feeds in their accounts. From this fact, it can be ascertained that as many as 93% of registered dowry MSME accounts have not been maximized in utilizing interactions with their followers. So, they need to understand the functions and benefits of engagement, between account owner and their follower in their Instagram account.

IV. CONCLUSION

Based on each stage of the research that has been carried out, several conclusions have been obtained, among others, this research produces an evaluation concept of the engagement rate based on followers, number of posts, likes and comments. The results of the assessment of 31 MSME accounts engaged in the delivery sector show that the majority of MSME actors, especially in the delivery sector, have not been able to optimize their business accounts in attracting their followers to participate in each post of the business account.

Engagement Rate is one of the most important benchmarks to find out the optimization of a business account on Instagram. By knowing the Engagement Level, we will know how much influence an account has on its followers. So that it can be used to find out and study deficiencies in terms of follower participation because the higher the engagement rate value means the better the performance and quality of an Instagram account.

For further research suggestions, a system is needed that is in the process of mining data using the Instagram API in order to measure the increase in followers from time to time. This will improve the quality of the measurement because it is known the change in followers based on a certain period of time. In this study, it is constrained that the account to be measured for the value of the engagement must be a public account so that there are several accounts that are set to private and the data cannot be retrieved.

REFERENCES

[1] A. P. J. I. I. APJII, "Laporan Survei Internet APJII 2019 – 2020 (Q2)," Asosiasi Penyelenggara Jasa Internet Indonesia (APJII), Jakarta, 2020.

[2] D. R. Indika and C. Jovita, "Media Sosial Instagram Sebagai Sarana Promosi Untuk Meningkatkan Minat Beli Konsumen," *Jurnal Bisnis Terapan,* vol. 1, no. 1, p. 25, 2017.

[3] A. A. Arman and A. P. Sidik, "Measurement of Engagement Rate in Instagram (Case Study: Instagram Indonesian Government Ministry and Institutions)," in *2019 International Conference on ICT for Smart Society (ICISS)*, Bandung, Indonesia, 2020.

[4] D. Henisa and N. Wilantika, "Content Characteristics of Government Social Media and The Impact on Citizen Engagement Rate," in *2018 Fourth International Conference on Advances in Computing, Communication & Automation (ICACCA)*, Jakarta, Indonesia, 2021.

[5] C. Primasiwi, M. I. Irawan and R. Ambarwati, "Key Performance Indicators for Influencer Marketing on Instagram," in *the 2nd International Conference on Business and Management of Technology (ICONBMT 2020)*, 2021.

[6] R. L. H. Yew, S. B. Suhaidi, P. Seewoochurn and V. K. Sevamalai, "Social Network Influencers' Engagement Rate Algorithm Using Instagram Data," in *2021 International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS)*, Jakarta, Indonesia, 2021.

[7] A. I. A. Rahim , M. I. Ibrahim, F. N. A. Salim and M. A. I. Ariffin, "Health Information Engagement Factors in Malaysia: A Content Analysis of Facebook Use by the Ministry of Health in 2016 and 2017," *International Journal of Environmental Research,* 2019.

[8] "Health in Digital Era 4.0: Analyzing Reader Engagement Rate on Instagram Account of Government Health Agencies," in *Proceedings of the International Conference on Health and Medical Sciences (AHMS 2020)*, Yogyakarta, 2020.

[9] C.-H. Liao, L.-X. Chen, J.-C. Yang and S.-M. Yuan, "A Photo Post Recommendation System Based on Topic Model for Improving Facebook Fan Page Engagement," *Symmetry,* 2020.

[10] H. Ryhänen, "Analysing Instagram Posts and Consumer Engagement," Arcada University of Applied Sciences, Helsinki, 2019.