

Entity Profiling to Identify Actor Involvement in Topics of Social Media Content

Puji Winar Cahyo^{*1}, Muhamad Habibi²

^{1,2}Department of Informatics, FTI UNJANI, Yogyakarta, Indonesia

e-mail: ^{*1}pwcahyo@gmail.com, ^{*1}muhammadhabibi17@gmail.com

Abstrak

Efisiensi dalam penggunaan media sosial mempengaruhi sifat dan cara komunikasi masyarakat modern, mereka lebih tertarik membicarakan sesuatu melalui media sosial daripada bertemu secara langsung di dunia nyata. Jumlah banyaknya pembicaraan pada suatu konten di media sosial bergantung pada topik yang dibicarakan, semakin menarik topik maka ketersediaan data di media sosial semakin meningkat. data tersebut dapat digunakan untuk melakukan analisis pengaruh aktor (akun mention) pada topik pembicaraan. Tingkat pengaruh suatu aktor dapat diukur dari berapa seringnya aktor tersebut dimention pada pembicaraan. Makalah ini bertujuan melakukan entity profiling pada konten media sosial untuk melakukan analisis pengaruh suatu aktor pada topik perbincangan. selanjutnya dengan menggunakan analisis sentimen dapat diketahui sentimen topik pembicaraan mengenai aktor tersebut. Metode Latent Dirichlet Allocation (LDA) digunakan untuk menganalisis dan memodelkan topik, sedangkan Support Vector Machine (SVM) digunakan untuk analisis sentimen. Hasil eksperimen menunjukkan bahwa pada topik dengan sentiment positive lebih mengarah pada keterlibatan akun penanggulangan bencana, sedangkan topik dengan sentiment negative lebih mengarah pada keterlibatan akun politisi, kritikus dan berita online.

Kata kunci— Entity Profiling, Pemodelan Topik, Analisis Sentimen, LDA, SVM

Abstract

The efficiency of using social media affected modern society's nature and communication; they are more interested in talking through social media than meeting in the real world. The number of talks on social media content depends on the topic being discussed. The more topic interesting will impact the amount of data on social media will be. The data can be analyzed to get the influence of actors (account mentions) on the conversation. The power of an actor can be measured from how often the actor is mentioned in the conversation. This paper aims to conduct entity profiling on social media content to analyze an actor's influence on discussion. Furthermore, using sentiment analysis can determine the sentiment about an actor from a conversation topic. The Latent Dirichlet Allocation (LDA) method is used for analyzes topic modeling, while the Support Vector Machine (SVM) is used for sentiment analysis. This research can show that topics with positive sentiment are more likely to be involved in disaster management accounts, while topics with negative sentiment are more towards involvement in politicians, critics, and online news.

Keywords— Entity Profiling, Topic Modeling, Sentiment Analysis, LDA, SVM

1. INTRODUCTION

Modern society makes social media an essential part of life. In addition to the harmful risks that arise, social media's benefits can support progress in several aspects. Such as the use of social media in the company's new product introduction survey to the public. The survey results are considered to be more efficient and give quite a positive impact. However, company members' active ability is needed to disseminate product information through social media [1]. Another use of social media is used to support the winning of the United States' presidential election. An analysis of public perceptions is a supporting parameter for increasing public interest in presidential candidates [2].

Efficiency in the use of social media affects the nature and way of communication of modern society. They are more interested in talking about something through social media than meeting in person in the real world. Not only that, but it is also supported by the freedom of society to be able to express or have an opinion. The amount of data on social media content depends on the attractiveness of a topic being discussed. The more interesting the issue being discussed, the availability of data on social media will increase. This can be used as a basis for analyzing the level of the actor's influence (account mention) on the topic of conversation. The level of influence can be measured by how often actors are discussed. To produce data on popular actor trends that can be searched for the sentiment of the conversation. These sentiments can be used as a basis for assessments to support decision making on popular issues that are currently being discussed.

Entity profiling is used to detect the relationship of actors to a topic being discussed. Entity profiling is worth researching, considering that social media is currently widely used to discuss hot issues that lead to a topic of conversation. Research on user profiling using temporal analysis based on tweet time intervals, it can be carried out to form clusters of recommended groups of hashtags and similar users to follow [3]. In the other case, clustering can be used to determine student profile base on learning achievement [4]. Not only that, entity profiling can be expanded to users' behavioral profiling with the similarity of tweet content attributes, post time, hashtag, and geolocation [5]. Also, Entity profiling can be done by linking account ownership detection with other open data, for example, wikidata [6]. The comparison of several methods that have been applied to this research is to use different stages of profiling. This research leads to collaboration between topic modeling, sentiment analysis, and entity search to form entity profiling of actors on the discussion topic's sentiment polarity.

This study focuses on conducting entity profiling on data on topics now being discussed on social media, Twitter. This research analyzes the sentiment on the tweet, then modeling the topic on the tweet. This process will produce topic segments that have positive and negative sentiment polarity values. The next step is to determine entity profiling to determine the actors involved in a conversation topic. So that specific topics can be identified as the value of the polarity of sentiment and its actors.

This study uses the Support Vector Machine (SVM) method to analyze sentiment regarding data on topics often discussed on Twitter social media. SVM method assigns values to selected words and phrases to create a text classification model [7]. SVM has a solid foundation and can perform classification with a higher accuracy level than other algorithms, especially high-dimensional data. One example is that SVM has the highest accuracy in domain text classification, as shown in the various research [8]. SVM is known as a method that has excellent accuracy values for classifying text data. One of the SVM applications for text classification is thesis title classification [9] and student comment classification [10].

Meanwhile, for modeling this research topic using the Latent Dirichlet Allocation (LDA) method. LDA is an unsupervised technique that automatically creates topics based on patterns of (co) occurrence of words in the documents that are analyzed [11]. Research related

to topic modeling using LDA includes modeling topics related to online hotel service reviews [12], modeling topics on scientific articles [13], modeling topics regarding road traffic [14], and modeling topics related to "ethnic marketing" at 239 journal articles published by nine major publishers [15].

2. METHODS

2.1 Tweet Data Extraction and Preprocessing

This research was conducted in several stages, including data extraction, preprocessing, sentiment analysis, topic modeling, topic analysis, entity profiling, and analyzing popular actors in a topic based on sentiment polarity. The stages of this research can be seen in Figure 1.

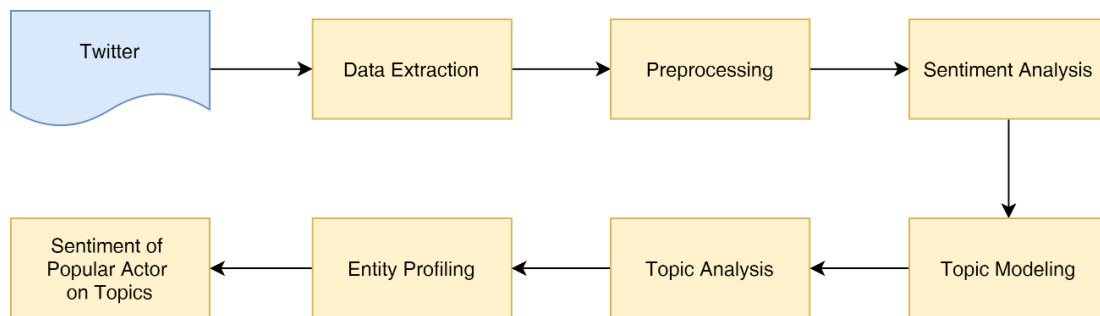


Figure 1 Research Stages

The data collected from 19 March to 13 July 2020. We managed to collect 263,125 tweets related to COVID-19. The data was obtained using the keyword "wabah corona". After the data has been collected, the next step is preprocessing. In general, the data tweets contain a lot of noise, such as meaningless words, misspelled words, including various abbreviations and slang words. These words often interfere with and reduce the performance of the resulting classification model. Therefore, tweets have to be preprocessed before actually extracting the features from them. The preprocessing steps that we take on the tweet data to be processed are:

- Tokenization: the process of dividing the text into specific parts.
- Normalization: brings the text to its standard form. The general normalization techniques used are as follows:
 - Case Folding: changes uppercase to lowercase
 - Elimination of periods in terms. - for example, M.C.S. to the MCS
 - Remove hyphens in a term. - for example, a medical-doctor to become a medical doctor.
- Cleaning: The steps in cleaning tweet data are as follows, remove the URL that is in the tweet, remove the hashtag (#) that is in the tweet, remove the number that is in the tweet, remove punctuation marks, such as question marks, exclamation points, periods, and others. And remove Unicode and symbols.
- Stopword Removal: delete deemed meaningless words using the stopword list.

2.2 Sentimen Analysis

Sentiment Analysis is the extraction of information that aims to obtain information about the author's feelings in positive or negative comments, questions, and requests by analyzing massive amounts of documents [16]. Sentiment analysis analyzes opinions, sentiments, evaluations, judgments, attitudes, and emotions of people towards product entities, administrative services, individuals, problems, events, topics, and attributes [17].

SVM is a Support Vector Machine. It is a non-probabilistic binary linear classifier. For a training set of points (x_i, y_i) , x is a feature vector, and y is the class. To determine the maximum margin hyperplane that divides the points with $x_i = -1$ and $x_i = 1$. The equation of the

hyperplane is; $w \cdot x + b = 0$. For a data set consisting of features set and labels set, an SVM classifier builds a model to predict the new examples' classes. It assigns a new case or data points to one of the categories [18].

Algorithm:

- a) Define an optimal hyperplane
- b) Extend step I for nonlinearly separable problems
- c) Map data to high dimensional space where it is easy to classify with linear decision surfaces.

2.3 Topic Modeling

Latent Dirichlet Allocation (LDA) is a generative probabilistic model for discrete data groups [19]. LDA is an unsupervised machine learning technique. This method aims to model documents that arise from various topics, where the topic is defined as the distribution of fixed word terms [20]. There are three generative processes for each document in the collection [21]. First, select a topic randomly from its distribution of topics for each document—the second step sample word distribution of words related to the chosen topic. Third, repeat the process for all words in the document. The visualization of LDA model representation can be seen in Figure 2 below.

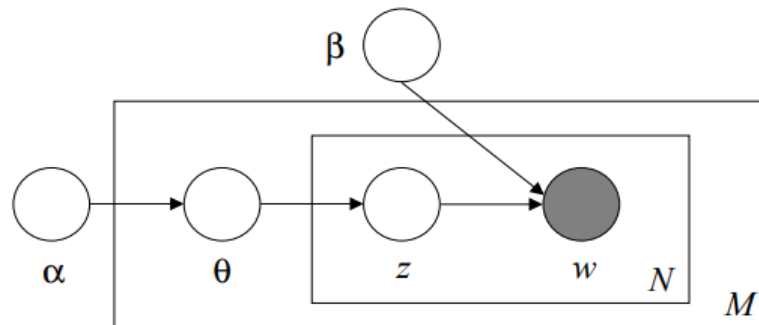


Figure 2 LDA model representation [19]

Figure 2 shows the three levels of LDA representation. The first level is corpus-level parameters, which are represented by symbol α and β . These corpus-level parameters are assumed to be sampled once in the process of generating a corpus. Secondly, document-level variables (θ) were tested once for each document—finally, word-level variables symbolized by z and w . Word-level variables are sampled once for each word in each document.

2.4 Entity Profiling

Entity profiling is extracting complete information about an entity using available data [22]. Extracting information on the Instagram platform is often necessary to determine how actor activities impact his followers [23]. In the Twitter platform, extracting information is generally carried out on actors' involvement on a conversation topic, which is generally marked with the hashtag # [24]. From the conversation topic, actors can be obtained by detecting named entities [25] or detecting mention @ in tweet content. Users can mention actors in a tweet and retweet due to several factors, including 1) influence is the actor's suitability factor with the context. 2) active users in using Twitter. 3) location of Twitter users [26]. This research seeks to detect actor sentiment by using mention @ on the topic of discussion. It is used to determine how many actors are talked in a positive or negative context.

3. RESULTS AND DISCUSSION

3.1 Sentiment Analysis

At this stage, we use the Support Vector Machine (SVM) method to classify sentiment polarity. We used 24,779 tweets for the training process, which will later create a classification model. The use of the SVM method for classification because it provides better accuracy results compared to other methods [27]. The evaluation methods used to see the SVM classification model's performance are precision, recall, accuracy, and F1-Score [28]. Evaluation of classifier performance is useful for seeing which classification method has a better performance value. The classification performance evaluation results show the amount of Precision 0.81, Recall 0.81, accuracy 0.81, and F-Score 0.81.

The sentiment polarity classification results from the COVID-19 tweet data show that the data is dominated by negative sentiment rather than positive sentiment. Tweets with negative sentiment had the number 257155, while tweets with positive sentiment had the number 5899. The results mean that the Indonesian people provide more negative opinions regarding the COVID-19 pandemic. Various things can trigger this negative opinion. It could be from government policies towards handling COVID-19 or even public complaints due to the COVID-19 pandemic.

3.2 Topic Analysis

In this experiment, we analyzed of the topic segment with a predetermined number of topics, namely 30 topics. This visualization is to show a physical representation of the word frequency distribution on each topic. After building topic segments with a predetermined range, evaluate the number of topic segments most suitable for further analysis. This study uses topic coherence to determine the ideal number of topic segments. Topic Coherence scored a topic by measuring the semantic similarity between words with a high score on a topic. This measure helps distinguish between topics that can be interpreted semantically and topics that are the result of human interpretation [29]. Topic Coherence is another way of evaluating topic models with much greater assurance of human performance. Figure 3 shows the results of measuring the amount of topic suitability using topic coherence.

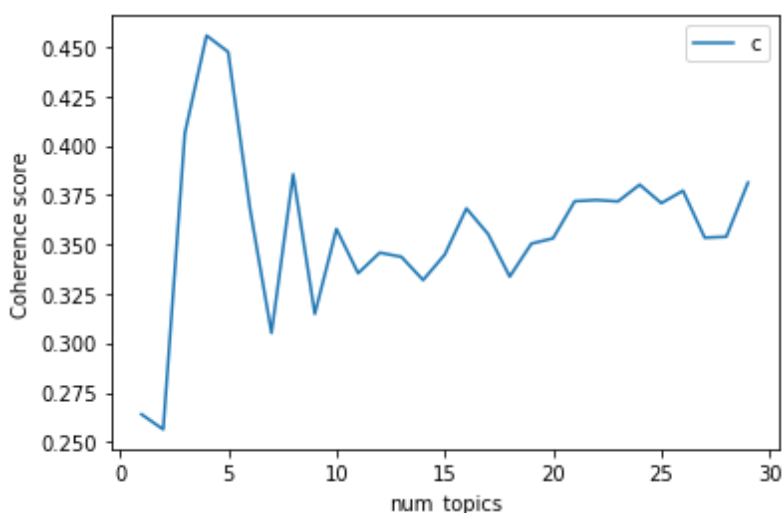


Figure 3 Coherence score of each number of topics

Based on Figure 3, it can be seen that the most suitable number of topic segments is the four topic segments with the highest topic coherence value, namely 0.45567. Therefore, the

subsequent analysis uses four topic segments. Figure 4 is a visualization of the distance map between topics generated for four topic segments.

Based on the frequency of words on each topic, analysis can be carried out related to the topic being discussed. Topic 1 is most concerned by the public regarding the COVID-19 pandemic, amounting to 33%. The topic discusses religious worship activities, in particular discussing preparation for activities in the month of Ramadan. Many Indonesians, especially Muslims, are restless because they cannot perform worship activities in mosques. The restlessness was due to the government's appeal to carry out Ramadan's worship activities at home [30].



Figure 4 Intertopic Distance Map Visualization

Topic 2 discusses the impact of Indonesia's economy during the COVID-19 pandemic, with a percentage of the number of tweets of 23%. The economic sector is the sector most affected by COVID-19, many large and small industries have closed, and many workers have been laid off. Indonesia's economic growth reached its lowest point in the second quarter of 2020. The government is expected to provide loans to small industries [31], which will absorb an additional workforce of 15 million people or 11.84 percent of the total workforce [32]. Furthermore, topic 3 discusses education during the COVID-19 pandemic, with a percentage of 21% tweets. Students have faced various problems related to anxiety, depression, poor internet connection, and unfavorable learning environments at home [33]. The issue of online learning is a problem that needs to be addressed by the government. Because so many students and parents complain about online learning, the learning process is not optimal.

The last topic, topic 4, discusses government policies during the COVID-19 pandemic with a percentage of 23% tweets. One of the successful handlings of COVID-19 is through procedures carried out by the government. The government needs to monitor and evaluate the COVID-19 handling protocol. This action is necessary because public perceptions are strongly influenced by the government's primary health care approach, impacting changing people's behavior [34]. The more people who care about the impact of COVID-19, the spread of COVID-19 will decrease. The data on the percentage distribution of tweets on each topic can be seen in Figure 5.

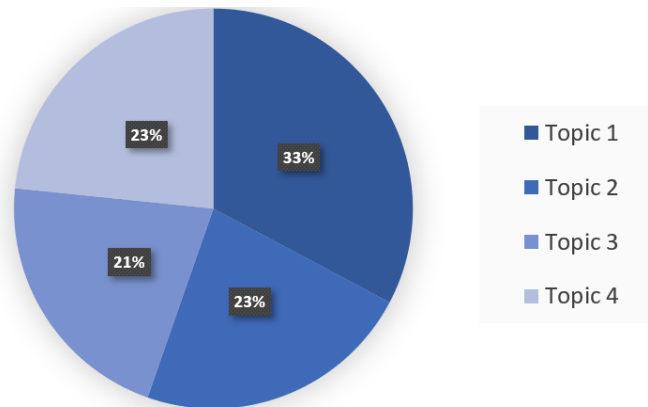


Figure 5 Distribution of tweets on each topic

3.3 Sentiment Analysis over Topic Category

This section outlines each topic's sentiments discovered in the previous area to complement the analytical perspective. Figure 6 shows the proportion of sentiment for each topic. We can see that all topics have a larger share of negative sentiment than positive sentiment. Topic 1, which discusses religious worship activities during the COVID-19 pandemic, has the highest negative tweets—then followed by topic 3, which discusses government policies in handling COVID-19 in second place, and topic 2 in the third, which discusses the economic impact of COVID-19. And in fourth place is topic 3, which discusses education during the COVID-19 pandemic. The results of these sentiments show that COVID-19 harms all aspects of human life so that most people give negative opinions in response to their complaints.

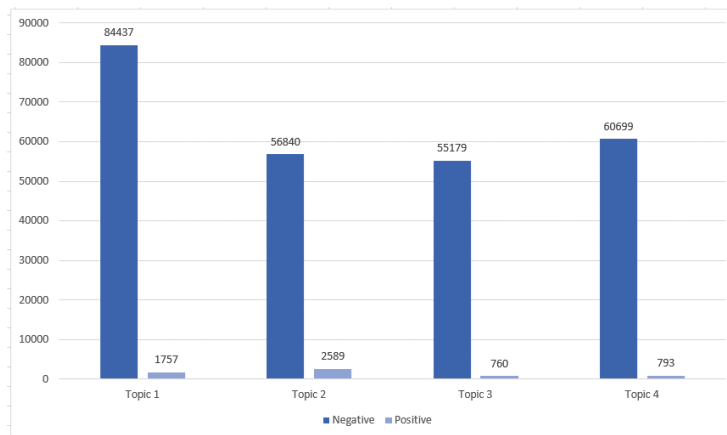


Figure 6 Sentiment proportion of each topic

3.4 Entity Profiling

The results of topic modeling from sentiment analysis show that overall topics are more likely to have a more significant negative sentiment than positive sentiment. We can be continued to entity profiling from topic modeling. It searches actors who are most discussed on each topic. The results of the top 10 discussions of actors are presented in Figure 7 to Figure 10.

It can be seen from Figure 7 that the community's response to @bnpb_indonesia's actions for handling the corona outbreak, especially in anticipation of religious activities, is considered quite positive, followed by the Director of Disaster Management Strategy Development himself, namely @aw3126. This was supported by a survey that showed that the Task Force's performance for the Acceleration of Handling COVID-19 received complimentary views from the public [35]. Meanwhile, negative talk about anticipation of religious activities

involving the @ginasnoer account is the highest, the second position, and so on involved a lot of online media including @detikcom, @bbcindonesia, and @cnnindonesia. Online media accounts have reported a lot about worship activities during the COVID-19 pandemic [36], [37], so that many tweets have led to negative sentiment.

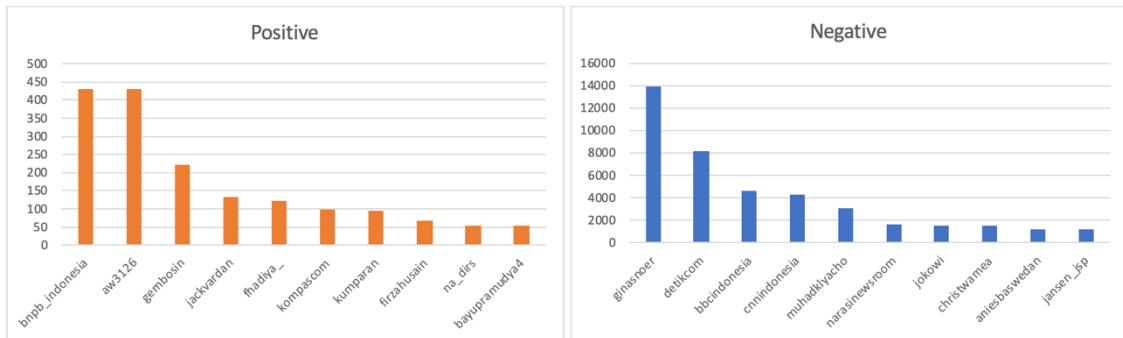


Figure 7 Popular Actor on Topic 1

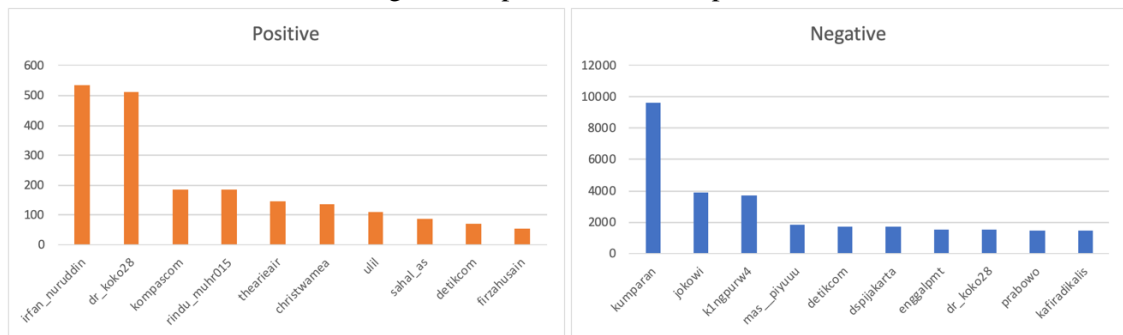


Figure 8 Popular Actor on Topic 2

As for Figure 8, it can be seen that community discussion is more likely to have an economic impact. Discussions with positive responses involved the most personal accounts, including @irfan_nuruddin, @dr_koko28, and @rindu_muhr015, while online news accounts involved @kompascom and @detikcom. Meanwhile, the negative talks involved accounts @kumparan, @jokowi, and @K1ngPurw4. As is known, the @K1ngPurw4 account is still suspended while the other accounts are still active. Accounts @kompascom, @detikcom, @kumparan issued news that highlighted Indonesia's economic conditions during the COVID-19 pandemic. Many tweets related to the economy mention the president's account @jokowi to convey people's aspirations.

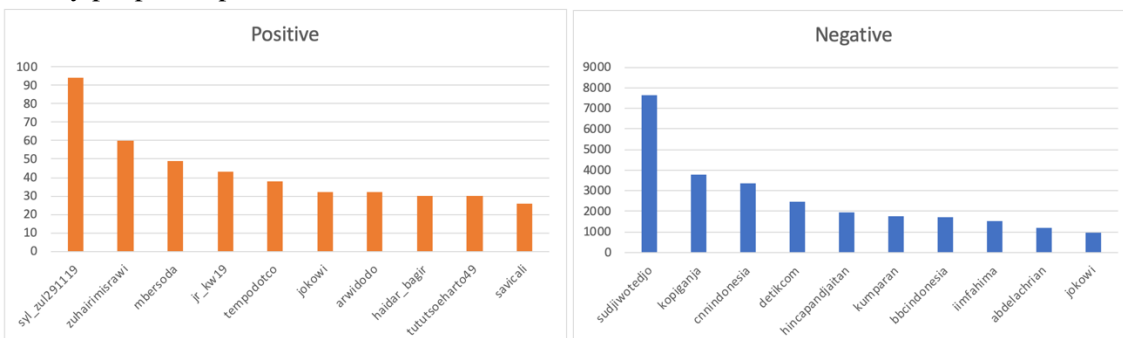


Figure 9 Popular Actor on Topic 3

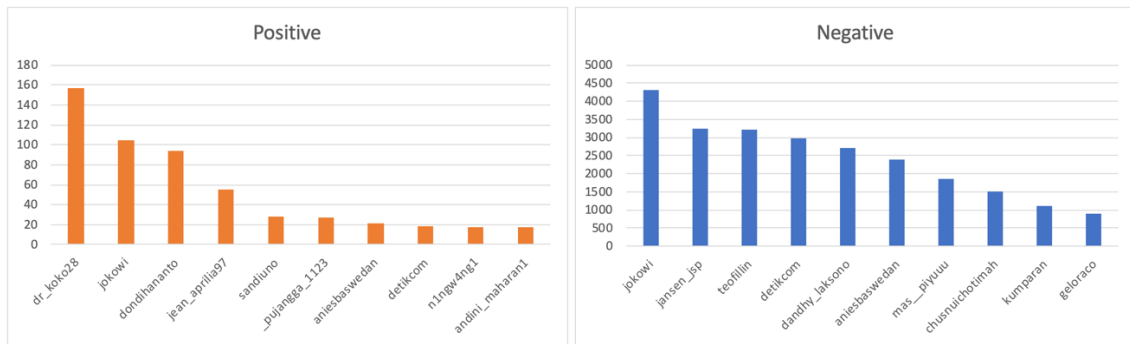


Figure 10 Popular Actor on Topic 4

Unlike the previous topic, in Figure 9, the community discussion about education during the pandemic period, the positive response involved more personal accounts with a discussion of no more than 100 mentions. Meanwhile, negative talk about education during the pandemic was more on individual tweets belonging to @sudjiwotedjo, @kopiganja, and @hincapandjaitan. Meanwhile, online news accounts include @cnnindonesia, @detikcom, @kumparan, and @bbcindonesia. One of the tweets from @sudjiwotedjo that went viral is that with the COVID-19 outbreak, parents who before the pandemic did not care about their child's education now can focus more on educating their children at home [38].

Figure 10 shows that there is only a little positive data from the government policy community's discussion regarding Indonesia's corona outbreak. This discussion involves several accounts, including @ dr_koko28, @jokowi, and @dondihananto. Meanwhile, the negative debate shows that there are quite many data, including the accounts @jokowi, @jansen_jsp, and @teofillin. The @jokowi and @jansen_jsp accounts are presidents and political figures, while the @teofillin account is a personal account that tweeted quite a lot of 124.2 thousand tweets. Discussions related to government policies regarding the corona outbreak in Indonesia involve many political members' accounts because political interests strongly influence government policies. Besides, the president's account @jokowi is included in the policy discussion because it has high policymaking authority.

4. CONCLUSIONS

This study succeeded in analyzing four topics from the tweet data with the keyword "wabah corona". From the topic modeling results, sentiment analysis was carried out to determine the polarity of sentiment on each topic discussed. The total number of topics generated has a more significant share of negative sentiment than positive sentiment. These results mean that Indonesians give more negative opinions regarding the COVID-19 pandemic. Various things can trigger this negative opinion; it could be from government policies towards handling COVID-19 or even public complaints due to the COVID-19 pandemic. The experimental results show that topics with positive sentiment are more likely to be involved in disaster management accounts, while topics with negative sentiment are more towards involvement in accounts of politicians, critics, and online news.

ACKNOWLEDGEMENTS

This research was conducted in 2020 of Penelitian Dosen Pemula (PDP) scheme funded by the Directorate of Research and Community Service Directorate General of Research and Development Strengthening (DRPM) of the Ministry of Research and Higher Education (Kemristekdikti) of the Republic of Indonesia. The research team would like to thank the

DRPM Kemristekdikti for allowing the research team to add insight and knowledge through research in this scheme. The data supported by Center of Study and Data Analytic Services of Universitas Jenderal Achmad Yani Yogyakarta. Hopefully, this research can bring benefits to the progress of the Indonesian nation

REFERENCES

- [1] D. Baum, M. Spann, J. Füller, and C. Thürndl, “Journal of Retailing and Consumer Services The impact of social media campaigns on the success of new product introductions,” *J. Retail. Consum. Serv.*, no. xxxx, pp. 0–1, 2018.
- [2] T. Macafee, B. Mclaughlin, and N. S. Rodriguez, “Winning on Social Media : Candidate Social-Mediated Communication and Voting During the 2016 US Presidential Election,” 2019.
- [3] A. Gorrab, F. Kboubi, A. Jaffal, and L. Grand, “Twitter User Profiling Model Based on Temporal Analysis of Hashtags and Social Interactions,” in *Natural Language Processing and Information Systems. NLDB 2017.*, 2017, vol. 1, pp. 124–130.
- [4] P. W. Cahyo, “Klasterisasi Tipe Pembelajar Sebagai Parameter Evaluasi Kualitas Pendidikan Di Perguruan Tinggi,” *Teknomatika*, vol. 11, no. 1, pp. 49–55, 2018.
- [5] S. PV and S. M. S. Bhanu, “UbCadet: detection of compromised accounts in twitter based on user behavioural profiling,” *Multimed. Tools Appl.*, vol. 79, no. 27, pp. 19349–19385, 2020.
- [6] M. Wiegmann, B. Stein, and B. Weimar, “Celebrity Profiling,” in *The 57th Annual Meeting of the Association for Computational Linguistics*, 2019, pp. 2611–2618.
- [7] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. United Kingdom: Cambridge University Press, 2000.
- [8] B. Liu, *Web Data Mining : Exploring Hyperlinks, Contents, and Usage Data*. Chicago: Springer, 2008.
- [9] A. F. Hidayatullah and M. R. Maarif, “Penerapan Text Mining dalam Klasifikasi Judul Skripsi,” in *Seminar Nasional Aplikasi Teknologi Informasi (SNATi) Agustus*, 2016, pp. 1907–5022.
- [10] M. Habibi and E. Winarko, “Klasifikasi Komentar Mahasiswa Menggunakan Kombinasi KNN berbasis Cosine Similarity dan Supervised Model,” no. x, pp. 1–11, 2017.
- [11] C. Jacobi, W. Van Atteveldt, and K. Welbers, “Quantitative analysis of large amounts of journalistic texts using topic modelling,” *Digit. Journal.*, vol. 4, no. 1, pp. 89–106, 2016.
- [12] Y. Guo, S. J. Barnes, and Q. Jia, “Mining meaning from online ratings and reviews: Tourist satisfaction analysis using latent dirichlet allocation,” *Tour. Manag.*, vol. 59, pp. 467–483, 2017.
- [13] H. Jelodar, Y. Wang, C. Yuan, X. Feng, X. Jiang, Y. Li, and L. Zhao, “Stabilization of an Inverted Robot Arm Using Neuro-Controller,” *Multimed. Tools Appl.*, vol. 78, pp. 183–198, 2018.
- [14] A. F. Hidayatullah and M. R. Ma’Arif, “Road traffic topic modeling on Twitter using latent dirichlet allocation,” in *2017 International Conference on Sustainable Information Engineering and Technology (SIET)*, 2017, pp. 47–52.
- [15] S. Moro, G. Pires, P. Rita, and P. Cortez, “A text mining and topic modelling perspective of ethnic marketing research,” *J. Bus. Res.*, vol. 103, pp. 275–285, 2019.
- [16] S. Mukherjee and P. Bhattacharyya, “Sentiment Analysis : A Literature Survey,” Indian Institute of Technology, Bombay, 2016.

- [17] B. Liu, *Sentiment Analysis and Opinion Mining*, no. May. Morgan & Claypool Publishers, 2012.
- [18] S. G. Kanakaraddi, A. K. Chikaraddi, K. C. Gull, and P. S. Hiremath, "Comparison Study of Sentiment Analysis of Tweets using Various Machine Learning Algorithms," in *Proceedings of the 5th International Conference on Inventive Computation Technologies, ICICT 2020*, 2020, pp. 287–292.
- [19] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet Allocation," *J. Mach. Learn. Res.*, vol. 3, pp. 993–1022, 2003.
- [20] D. M. Blei and J. D. Lafferty, "Topic Models," in *Text Mining: Classification, Clustering, and Applications*, 2009, p. 71.
- [21] L. Hong and B. D. Davison, "Empirical study of topic modeling in Twitter," in *Proceedings of the First Workshop on Social Media Analytics - SOMA '10*, 2010, pp. 80–88.
- [22] S. Varma, N. Sameer, and C. R. Chowdary, "ReLiC: entity profiling using random forest and trustworthiness of a source," *Sādhanā*, vol. 44, no. 9, p. 200, 2019.
- [23] P. W. Cahyo and M. Habibi, "Clustering followers of influencers accounts based on likes and comments on Instagram Platform," *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 14, no. 2, pp. 199–208, 2020.
- [24] M. Habibi and P. W. Cahyo, "Clustering User Characteristics Based on the influence of Hashtags on the Instagram Platform," *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 13, no. 4, pp. 399–408, 2019.
- [25] P. W. Cahyo and E. Winarko, "Model Monitoring Sebaran Penyakit Demam Berdarah di Indonesia Berdasarkan Analisis Pesan Twitter," Universitas Gadjah Mada Yogyakarta, 2017.
- [26] Y. Li, Z. Ding, X. Zhang, and B. Liu, "Confirmatory Analysis on Influencing Factors When Mention Users in Twitter," in *2nd International Workshop on Web Data Mining and Applications (WDMA 2016)*, 2016, vol. 1, pp. 112–121.
- [27] M. Habibi and P. W. Cahyo, "Journal Classification Based on Abstract Using Cosine Similarity and Support Vector Machine," vol. 4, no. 3, pp. 48–55, 2020.
- [28] A. A. Anees, H. Prakash Gupta, A. P. Dalvi, S. Gopinath, and B. R. Mohan, "Performance analysis of multiple classifiers using different term weighting schemes for sentiment analysis," in *2019 International Conference on Intelligent Computing and Control Systems, ICCS 2019*, 2019, pp. 637–641.
- [29] K. Stevens, P. Kegelmeyer, D. Andrzejewski, and D. Buttler, "Exploring topic coherence over many models and many topics," in *Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP)*, 2012, no. July, pp. 952–961.
- [30] F. WH, "Pedoman Ibadah Ramadan dan Lebaran di Masa Pandemi Covid-19 | Indonesia.go.id," *Indonesia.Go.Id*, Apr-2020. .
- [31] A. W. Bartik, M. Bertrand, Z. Cullen, E. L. Glaeser, M. Luca, and C. Stanton, "The impact of COVID-19 on small business outcomes and expectations.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 117, no. 30, pp. 17656–17666, Jul. 2020.
- [32] N. Wira Sakti, "Perekonomian Indonesia Pasca-Pandemi Covid-19 Halaman all - Kompas.com," *Kompas.com*, May-2020. .
- [33] N. Kapasia, P. Paul, A. Roy, J. Saha, A. Zaveri, R. Mallick, B. Barman, P. Das, and P. Chouhan, "Impact of lockdown on learning status of undergraduate and postgraduate students during COVID-19 pandemic in West Bengal, India," *Child. Youth Serv. Rev.*, vol. 116, p. 105194, Sep. 2020.
- [34] R. Jose, M. Narendran, A. Bindu, N. Beevi, M. L, and P. V. Benny, "Public perception

- and preparedness for the pandemic COVID 19: A Health Belief Model approach,” *Clin. Epidemiol. Glob. Heal.*, Jun. 2020.
- [35] T. K. Yudianto, “Survei: 63% Responden Puas Atas Kerja Gugus Tugas Tangani Covid-19 - Nasional Katadata.co.id,” <https://katadata.co.id/>, Jun-2020. .
- [36] C. Indonesia, “Panduan Lengkap Pembukaan Rumah Ibadah di Masa Pandemi,” <https://www.cnnindonesia.com/>, May-2020.
- [37] P. Yasmin, “Panduan Ibadah di Masjid Selama Pandemi dari MUI DKI Jakarta,” <https://news.detik.com/>, Jun-2020.
- [38] R. Gunadha, “Budayawan Sudjiwo Tedjo Beri Ucapan Terima Kasih pada Corona, Mengapa?,” Sep-2020.