Buzzer Detection to Maintain Information Neutrality in 2019 Indonesia Presidential Election

Adrian Julianto Panatra[1], Felisitus Bryan Chandra[2], Wiliam Darmawan[3]
Computer Science Department ,
School Of Computer Science
Bina Nusantara University
Jakarta, Indonesia 11480
adrian.panatra@binus.ac.id[1],
felisitus.chandra@binus.ac.id[2],
wiliam.darmawan@binus.ac.id[3]

Harco Leslie Hendric Spits Warnars
Computer Science Department, BINUS
Graduate Program - Doctor of Computer Science
Bina Nusantara University
Jakarta, Indonesia 11480
spits.hendric@binus.ac.id

Wiranto Herry Utomo
Faculty of Computing
President University
Jawa Barat, Indonesia 17550
wiranto.herry@president.ac.id

Tokuro Matsuo
Graduate School of Industrial Technology
Advanced Institute of Industrial Technology
Tokyo, Japan
matsuo@tokuro.net

***Abstract*—This paper proposed a method which detects a political buzzer in social media, specifically Instagram. With Indonesia undergoing 2019 presidential election, a detection of buzzers that causes much trouble in maintaining information neutrality is seen as a needed. One of the many reasons is because those buzzers spread false news making the information gained by the use of social media to be not neutral and deliberately offends or attack those that they are not in favor of. Those buzzers share a similar characteristic, tendency, or even possess the same pattern. Grouping, classification, and detection method are used to counter this problem. This research gives a slight overview of what is happening in social media and a theory of how to deal with those problems. The argument is expected to help to identify buzzer in real life, thus helps in maintaining information neutrality along with the social media in Indonesia.**

Keywords*—*buzzer detection, 2019 Indonesia Presidential Election, social media retrieval

I. INTRODUCTION

According to *Asosiasi Penyelenggara Jasa Internet Indonesia (APJII)*, 143 out of 262 million Indonesia citizen (54.68%) is connected to the internet by 2017. Among those 143 million citizens, 87.13% of them are accessing social media as their lifestyle [1]. Social media is an efficient yet effective medium in delivering information to others. Social Media as a media with high social dynamic and allows an open communication to many parties with various agenda is the right medium to amplify social participation [2]. It also gives a benefit in building personal branding as it does not use any trick or temporary popularity but the audience assessment. Several social media acts as a media to communicate, discuss, and even in bringing reputation [3]. Where buzzers come in to gain attention in the news or personal branding and this discussion, the buzzer meant is the anonymous account used to fulfill a specific political agenda.

Social media (buzzer) has been playing a critical role in recent Indonesia elections (2012 Jakarta Governor's election[4], 2014 Presidential election[5], and 2017 Jakarta Governor election[6]). According to Howard & Hussain [7], this happens because of there something new that's been given to the activist regarding freedom: difficult to control information network for the state and coordination tools installed in trusted systems close relatives and colleague circle. This causes buzzers to try to make benefit out of the situation creating a popularity increases of the political party they are supporting or even violating some regulation to strike down their opponent. These buzzers usually use trickery or spread some fake news, especially in a political campaign period. They desperately defend their political agenda argument by offending their opponent's case with any way possible. Dense debates happening between these political buzzers are polluting social media, mainly containing anger, rants, and offensive statements [8]. For example in 2014's presidential election, several issues such as *hidden fear (Jokowi was said as a Communist, non-muslim, a Chinese, and even meant to be dead, whereas Prabowo was accused of being a NAZI, authoritarian, New Order's leadership style, the oppressor of the weak). Understandable folks appeal (Jokowi is a modest man, and Prabowo is a conglomerate), testimony, and Name Calling (Jokowi is a puppet of Megawati)* are thrown at both candidates by buzzers account, especially the last eight weeks before the election.[5]

Moreover, through the end of 2018 to 2019, Indonesia will undergo yet another political campaign period for almost a year as the presidential election will occur on 17th April 2019. Such a historical event is sure to be a hot and trending topic to be discussed by the Indonesian social media. Consider several conditions stated in the previous paragraph; this paper aims to detect the buzzers to help to maintain information given in Indonesia's social media to be fair and neutral.

II. RELATED WORKS

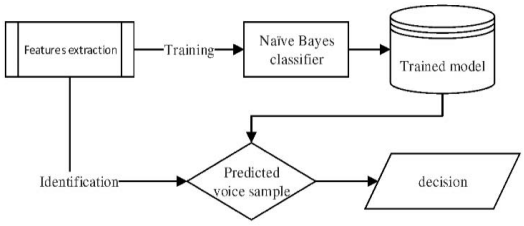


Figure 1. Naive Bayes 1

978-1-7281-2627-2/19/$31.00 ©2019 IEEE
DOI 10.1109/IIAI-AAI.2019.00177

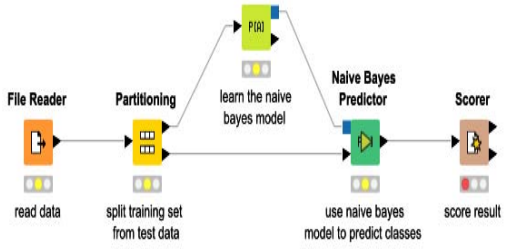


Figure 2. Naive Bayes 2

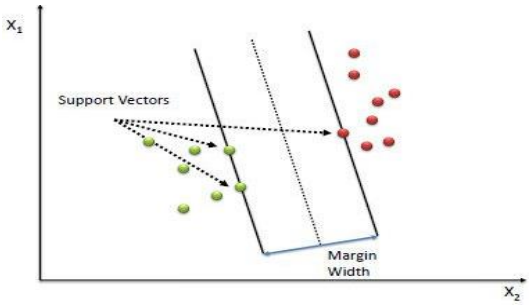


Figure 3. Support Vector Machine 1

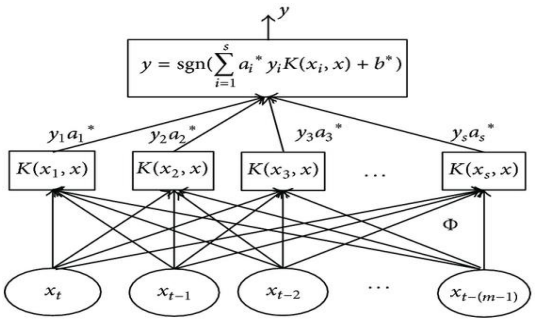


Figure 4. Support Vector Machine 2

Before identifying a buzzer, data classification is needed. There are a set of rules that can be defined to help the decision of what features to be classified later. Set of rules that help to give an insight into whether an account has a human, bot, or fake account behavior. Moreover, Camisani-Calzolari[19], Van den Beld[20], and Socialbakers[20] set of rules methods can be used as an example. Those methods were implemented on Twitter, but based on them, the rules set of Instagram accounts could be made.

In this article, the data is first classified what kind of group it is. This classification is placed depending on the features of the data. Several works have tried to detect buzzer using different ways. Methods like Naive Bayes, Support Vector Machines (SVM), random forest, and feature-based detection were used to identify buzzers. Most of this buzzer detection was implemented on several Online Social Networks (OSN's), such as Facebook, Twitter, RenRen, Tuenti, etc., but none of them already perform it on Instagram. In this section, those methods are reviewed according to its relation with this article.

Naive Bayes is a simple technique used to construct a classifier, like a model, by using assumption among the features. Models are represented in the form of feature value's vector. This method is commonly used to detect spams and was one of the first method used to identify spam in earlier years. Spam is one of the features related to buzzers. Thus we use Naive Bayes to help us in detecting buzzers. [9] Figure 1 and Figure 2 describe the typical architecture of Naïve Bayes.

Support vector machines are supervised learning models used to analyze data in classification or regression analysis [10]. As seen in Figure 3, it makes a model out of two different categories, checks if the new sample is categorized as a category or the other, and represented as points in space and divided by a gap (the more significant difference is better). [11] Cao et al. [12] use the SVM algorithm to look for an optimal hyperplane as a decision function in a high-dimensional space. One of the reason is that they needed a non-linear classifier and SVM with radial basis function (RBF) kernel. Figure 4 helps to describe the calculation used in SVM.

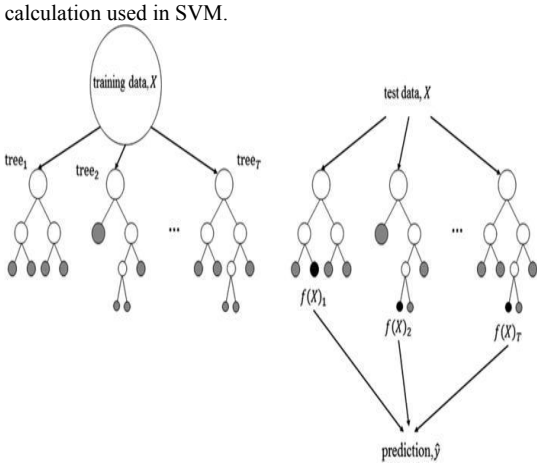


Figure 5. Random Forest

Random forest used to combines all data classifier, like a decision tree, into the better version of the classifier. The way of changing the first sample of the classifier is replacing from the original training set to the new training set with the same size. Then for each node in the decision tree, we choose the random *m* features and split the decision tree according to the most right possible split among all the *m* features. Value of *m* is essential to be chosen for balancing the strength of individual trees. When adding a new sample, the resulting model scores get from running the example through all the trees; after that, the result from it will be combined. Because our case is a binary classification, then the score is a percentage of trees that give a positive result with the sample that used [12] and this is illustrated in

Human Behaviour Rulesets
1. Profile name without a number
2. The image is not default
3. Has ever uploaded Insta Story
4. Has post's archive
5. Posts more than 10
6. Connected with Facebook Account
7. Use punctuation or emoticon in post's caption

Bot Behaviour Rulesets
1. Five or more post in a day
2. The ratio followers/likes of the account under the expected 50:1, or less
3. More than 30% of all the post of the account use hashtag, such as "#jokowi," "#prabowo," or "#2019gantipresiden."
4. The same posts are repeated more than three times, even when it posted to the different account
5. More than 90% of the account posts are reposts
6. The account has never post
7. The account is more than two months but still use a default profile image
8. The user-following more than 100 accounts without gains followers

Figure 5.

Figure 6. Rule Sets

Feature-Based detection detects different user activities. It is applied by gathering evidence(s) from the features[13]. As an example, Instagram uses the post feature as a vessel to gather evidence. Google features an uploaded photo or picture to use as its search material where to find a similar or exact picture. It needs to gather probabilities from each angle of photos and similar object spot that appears in the photo. Features are then classified to help the learning process later. Figure 7 helps in illustrating those.

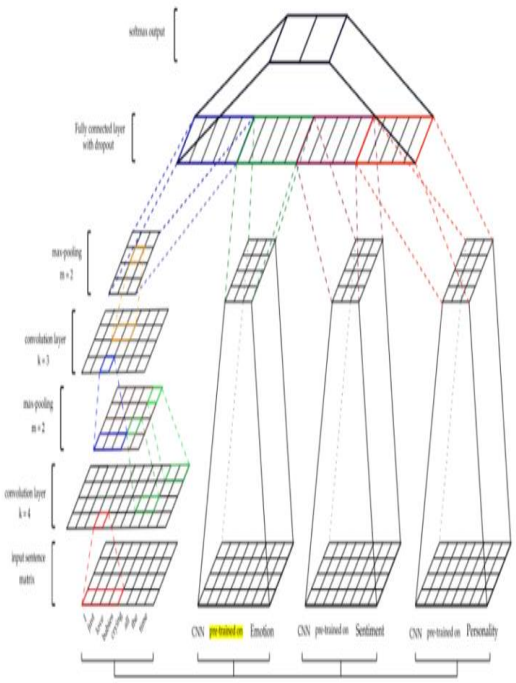


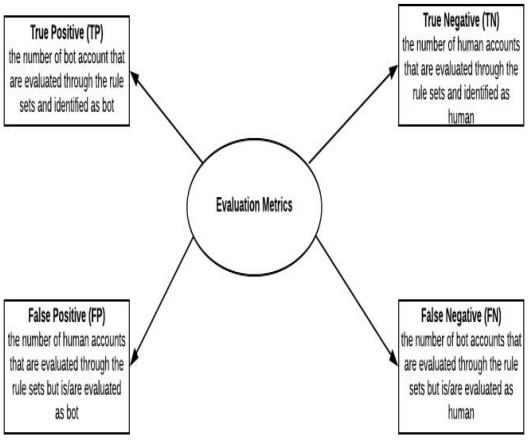
Figure 7. Feature-Based Detection

III. PROPOSED METHOD

Based on reviewed methods for classification and detection explained in the previous section, we would like to suggest some way in detecting buzzers on Instagram. Before doing any process of classification nor discovery, datasets are needed to be taken first [14,15]. Several companies provide online tools to classify Instagram followers based on their fakeness degree. Now, we use HypeAuditor to detect the quality of the audience and also the interest of the account audience. Datasets are then grouped based on their interest. The fakeness of the account followers increases on several criteria listed in Figure 6.

Before getting to the features, rules set are made to help in defining characteristics that are going to be used to detect the buzzer. As Camisani-Calzolari, Van den Beld, and Socialbakers set of rules methods are used on Twitter, this article would like to propose a set of rules too, particularly in Instagram. The set of rules are made based on those method used before to see if an account is considered as a human, bot, or fake account. Set of rules that are made to identify them are illustrated in Figure 6.

Those rule sets are next evaluated if it has considered as a human or bot. The experiment should be done to evaluate them and be summarized to several following evaluation



metrics called True or False Negative or Positive. Those evaluations descriptions are illustrated in Figure 8 below.

Figure 8. Evaluation Metrics

Buzzers can be deducted from the features such as post's time, images, a frequency of posts, hashtags, and ordinary followed account [16]. Buzzers tend to have a high rate of a post with seemingly same pictures and a little to no difference of post time (patterns). [17] They also tend to have a commonly followed account and same hashtag. An account that fits those criteria can be classified as Buzzers. In this particular article, **political buzzers** are of our primary concern [18], and in that case, the hashtag feature would be narrowed into a mainly political hashtag, for example, #pemilu2019, #pilpres2019, #jokowi, #prabowosubianto, #2019GantiPresiden, #caprescawapres2019, and so on. When the grouping is finished, classification is done to those datasets.

Classifications are done by using Naive Bayes, Support Vector Machines, and Random Forest algorithms. In this model, we consider the result that's related to the Buzzer definition. We count the probability of the post that user post will occur at the result related to the Buzzer definition. Now we will consider features such as #pemilu2019, #pilpres2019, #jokowi, #prabowosubianto, #2019GantiPresiden, #caprescawapres2019 using naive Bayes classifier. Assume that we have a set of posting that occurs with the following hashtag that already told before. Let us assume that we also know, using the training, about the behavior of the person that post about that. Also, knowing these attributes, we find out if any other account that posts the same post occurred. The other attributes that we considered is an account that already declared to be a buzzer. The naive Bayes classifier works by the probability of one attributes and does not affect the other probability.

Moreover, by using SVM, the data are analyzed and put based on their categories. Data are maps into dimensional space to maximize the gap between point. The wider the differences, the more optimal the decision function is. The

classifier is then improved later on by using Random Forest. All those data classifier are combined and are replaced through training that split the decision tree according to the most right possible splits without changing the features.

Finishing the classification, data are detected if it is a buzzer or not by using feature-based detection. Buzzers are usually more than one account being controlled by the same person; this makes it have a similar tendency. Features are made into numbers and labels (Buzzer or Normal Account) to identify them. High frequency of posting content, little to the same time of posting, similar images or hashtags in content, and following an ordinary account are features that can be said as Buzzer(s). Figure 9 shows the process diagram of overall detection.

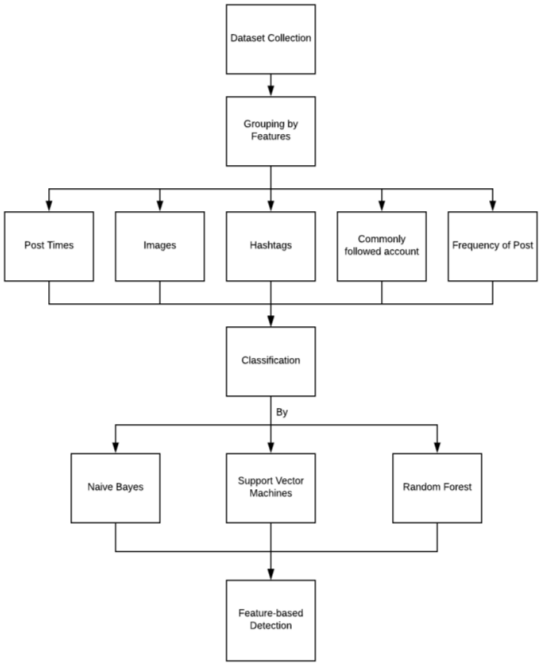


Figure 9. Proposed Method Process Diagram

IV. CONCLUSION AND FUTURE WORK

In this research, buzzer, especially the political one, is the primary target to be detected. A method is proposed to achieve this result. As explained in the previous figure, firstly dataset is gathered through Instagram Graph API. Features are then grouped based on their post times, images, hashtags, commonly followed account, and frequency of post. These features are then classified by using Naive Bayes, Support Vector Machine, and Random Forest method. These methods are the most commonly used to detect fake accounts. Buzzers are also included as counterfeit accounts. Thus, the methods are used. As there are only limited resources, this research is solely to propose a theory to deal with buzzer by using several methods. An argument that is to help and give an overview of the future researcher in experimenting with the same matter. In the future, improvement is expected. It is hoped that after the detection succeeds, those accounts can be tracked through their IP and the irresponsible person behind it can be apprehended. By doing this, a better and comfortable social media environment can be achieved, and information neutrality is maintained.

REFERENCES

[1] Indonesia Internet Service Provider Association. "Survey Penetrasi & Perilaku Pengguna Internet Indonesia." Internet: https://apjii.or.id/survei2017/download/WF1lRD24ZzBHpxemXGKgrV6JCvjQcw, 2017. [Oct. 1, 2018]

[2] D.R. Rahadi. "Perilaku Pengguna dan Informasi HOAX di Media Sosial". *JURNAL MANAJEMEN & KEWIRAUSAHAN*, vol.5, no.1, pp. 58-70. 2017.

[3] Junaidi, F. Alfiah, E. Susanti, J. Kristinna, O.R. Ardiansyah, D.Pradipta, and Wulaningsih. "Manfaat Menganalisis Pengaruh Sosial Media Facebook Terhadap Kampanye Partai Politik Di Indonesia". SEMNASTEKNOMEDIA, Feb. 2015.

[4] W.P. Utomo. "Menimbang Media Sosial dalam Marketing Politik di Indonesia: Belajar dari Jokowi-Ahok di Pilkada DKI Jakarta 2012". Jurnal Ilmu Sosial dan Ilmu Politik, vol. 17, pp. 67-84. Jul. 2013

[5] N. Hasfi, H.P. Santosa, and T. Lukmantoro. "Twitter dan Pemilu di Indonesia: Propaganda di Ruang Publik Maya" in *Information and Communication Technology, dan Literasi Media Digital,* 1st ed. I.P. Hadi, Ed. Surabaya: ASPIKOM, 2015, pp. 23-36 .

[6] B.A. Amsak, and U. Salamah. "Post Truth and Challenger's Issues Management Via Twitter in Jakarta's Governor Election 2017" in Proceeding of Indo-IGCC. 2017, pp. 320-336.

[7] L.A. Abdillah. "Social Media As Political Party Campaign In Indonesia". Jurnal Ilmiah MATRIK, vol.16, pp. 1-10. April. 2014.

[8] I. Syahputra. "Demokrasi Virtual dan Perang Siber di Media Sosial: Persepektif Netizen Indonesia". *JURNAL ASPIKOM*, vol. 3, pp. 457-475. Juli. 2017.

[9] M. Granik and V. Mesyura, "Fake news detection using naive Bayes classifier" in Proceeding of IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON), pp. 900-903, 2017.

[10] W. Contributors. "Support vector machine" Internet: https://en.wikipedia.org/wiki/Support_vector_machine, Oct. 14, 2018. [Oct. 15, 2018]

[11] A. H. Marti, "Support Vector Machines," *IEEE Intelligent Systems*, 1998, pp. 18-28.

[12] C. Xiao, D.M. Freeman, and T. Hwa. "Detecting Clusters of Fake Accounts in Online Social Networks" in Proceedings of the 8th ACM Workshop on Artificial Intelligence and Security. 2015, pp. 91-101.

[13] Y. Boshmaf, D. Logothetis, G.Siganos, J. Leria, J. Lorenzo, M. Ripeanu, and K. Beznosov. "Íntegro: Leveraging Victim Prediction for Robust Fake Account Detection in OSNs". *Computers & Security*, vol. 61, pp. 142-168. August. 2016.

[14] Mohammadrezaei, M., Shiri, M. E., and Rahmani, A. M, "Identifying Fake Accounts on Social Networks Based on Graph Analysis and Classification Algorithms," *Security and Communication Networks*, vol. 2018, pp. 1-8. August. 2018.

[15] A. I. Rodrigo, B. J. Sylvio, C. S. P. Kátia, S. K. Guilherme, C. G. Rodrigo, L. P. J. Mario, and N. D. S. Ivan, "Account classification in online social networks with LBCA and wavelets" *Information Sciences,* vol. 332, pp 72-83, November 2015. [Online]. Available: www.sciencedirect.com. [Accessed Oct. 18, 2018].

[16] J. Ratkiewicz, M. D. Conover, M. Meiss, B. Goncalves, A. Flammini, and F. Menczer, "Detecting and Tracking Political Abuse in Social Media," in *Complex Networks and Systems Research: Proc. of the 5th Int. AAAI Conf. on Webblogs and Social Media*. 297-304.

[17] G. Supraja, S. W. Joshua, H, Brian, and N. M. Jeanna, "Profile characteristics obtained using an activity-based pattern detection approach," in *Fake Twitter Accounts: Proc. of Int. Conf. on Social Media and Society*, 2015.

[18] M. Ibrahim, O. Abdillah, A. Wicaksono and M. Adriani, "Buzzer Detection and Sentiment Analysis for Predicting Presidential Election Results in a Twitter Nation" in Proceeding of IEEE International Conference on Data Mining Workshop (ICDMW), pp. 1348-1353, 2015.

[19] M.Camisani-Calzolari, (digitalevaluations), Analysis of Twitter Followers of the US Presidential Election candidates: Barack Obama and Mitt Romney, August 2012.

[20] S. Cresci, R.D. Pietro, M. Petrocchi, A. Spognardi, and M. Tesconi, "Fame for sale: Efficient detection of fake Twitter followers" in *Decision Support Systems*, vol. 80, pp 56-71. December. 2015