Emotion Mining of Indonesia Presidential Political Campaign 2019 using Twitter Data

A A Arifiyanti¹, E D Wahyuni², and A Kurniawan³

¹, ², ³ Information System Department, University of Pembangunan Nasional “Veteran” Jawa Timur, Indonesia

Coresponding author’s email address: ¹amalia_anjani.fik@upnjatim.ac.id

The presidential election campaign in Indonesia which was held in 2019 attracted the interest of the Indonesian people. People expressing their emotion through social media, one of which is Twitter. Supervised learning is used to classify user’s emotion of that issues. The goal of this research is extracting the emotional content in tweets. The interest is in whether the text is an expression of Twitter users’ about the presidential candidates and whether this online forum represents the election results. For this purpose, text mining techniques are performed and we will compare Naive Bayes Classifier and SVM in classification process. The result showed that Naive Bayes Classifier and SVM have good performance for classifying text and each of classifier have not outperform the others. The result also showed that stemming step in text pre-processing do not give any significant result in model accuracy, because twitter data use informal language like abbreviation or slang.

Keywords : Elections, Emotion Mining, Text Classification, Twitter

1. Introduction
Indonesia presidential election is huge event which attract Indonesia people attention. Indonesia people have different emotion about presidential candidates. Their opinion about the candidates expressed in many ways, one of them is social media. Through 280 characters limit on social media Twitter, Indonesia people expressed their emotion about this campaign with. In this presidential election, there are four debates of the presidential candidates to convey their vision in five years they will govern. After each debates, there are many hashtag trending in Indonesia about this debate such us #pilpres2019 and many more [1]. Indonesia people emotion about each presidential candidate and overall presidential election need to be analyze. These emotions will affect image of each presidential candidate.

Detecting people emotion can be done by mining through their writing in twitter. Twitter is a social media platform which commonly used by people in expressing their emotion about such issues or events. Presidential election can be predicted from people’s writing in their twitter [2] [3]. Many researches detect emotion based on their sentiment (positive or negative or neutral sentiment) [4], they argued these sentiment are enough for their analysis and no need to dig deeper. But for some cases, this approach has limitation which sentiment can not detect specific emotion such as sad or joy. So, some research approach does not limit in positive or negative sentiment but more specific emotions. These approached used a more sophisticated solution that digs far past the simple negative and positive, instead looks to analyze specific emotions and how they lead to future behaviors. A set basic emotion in recent research, with four emotions (sad, joy, anger, and surprise), used in this research. The important thing is sentiment or emotion classification about dynamic event have many challenges because of it’s dynamic data [5].
Classifying emotion into emotions can be used with many approach. Unsupervised and supervised method can be used for classifying people emotion into binary-class or multi-class. Classifier such as naive bayes or SVM are commonly used to predict people emotion about political campaign because these classifier have good performance toward text. While some research showed that emotion mining can predict the election [6], but some cases showed opposite result [7]. Even though some of these researches have different result, but this kind method can help for seeking potential information for enhancing our knowledge of political condition [8].

In this study, we focus only on text twitter which contain four basic emotion. We will use standart text pre-processing for extracting the feature. These data will be classify with text mining process and classify into four classes with supervised classifier. We will try naive bayes classifier and SVM which commonly used for text classification and compare the result.

2. Emotion and Emotion Mining

Human have distict emotion which expressed through their face or body language. It’s triggered by different issues, events, or other interests. Until now, researchers does not have agreement about human’s basic emotion. Generally emotion have two polarity (positive and negative). The earliest research, Ekman model, suggest there are six different basic emotion [9]. Later, there are different kind of emotion arranged emotion in polarity axes or opposite of each other like joy vs sadness [10]. While there are some emotion that have opposition, but there are some emotion which have no clear distiction over polarity axis and can be both of them such as surprise depending on the situation. There are anger, fear, joy, and surprise are common in some emotion models [11]. While there are no agreement in set basic emotion, the earliest emotion model suggested, with six emotions, is the most widely used. Recently there is research suggested four basic emotion [12] rather than six, opposed the Ekman emotion model. These emotion are sad, joy, anger, and surprise. This research argued anger-disgust have same signal at forming expression in human facial. This also applies in fear-surprise too.

People emotion can affect how they react about something such as events or issues. These emotion can be detect or classify through a mining technique. Emotion mining is a method for analyzing people's emotions about something or any interest. Text emotion mining is a branch topic of emotion mining based on their writings. Text emotion mining can be done with classification process [13]. The text need to be pre-processing first for extracting and selecting important features. This process will affect the evaluation model, so many studies suggested different combination of pre-processing steps [14] [15] [16]. After this process, classification process will be done with classifier algorithm such as Naive Bayes Classifier [17] [18] [19] or SVM [20] [21]. These classifiers are commonly used because they have good performance toward text classification.

3. Dataset

In this research we scraped public tweets about Indonesia presidential election between February and March 2019. We collected all tweets that contained the names of two president candidates and keyword hashtag about presidential election (#pilpres2019). The scrapped data then divided into two part. First part is tweets which expressed emotion and second part is neutral tweet. We use the first part to be mined for classifying emotion about presidential election.

We use 397 tweets which is devided into four classes. These classes based on four kind of emotions which is 1) joy, 2) sad, 3) anger, and 4) surprise. The labelling process based on subjective judgment for labelling the tweets into their emotion classes. This will be our training data.

| Emotion class | Number of tweets | Tweet example |
|---------------|------------------|---------------|
| Joy           | 98               | Mantaaap pokoknya #DenganMusikKitaSatu Dengan @Jokowi #01IndonesiaMaju |
| Sad           | 100              | RT @diah_soewarno: Kenapa pak Prabowo sering |
| Emotion class | Number of tweets | Tweet example |
|---------------|------------------|---------------|
| Anger         | 100              | @faizalassegaf @jokowi Ada prestasi kau claim, Bgitu ada kesalahan kau suruh tanya ke pejabat atau instasinya alias gak mau tau..δŸ—•δŸ•¼ |
| Surprise      | 99               | RT @03_Nakula: FBR dulu dukung anis-sandi di pilkada DKI dalam pilpres 2019 mereka memberikan dukungan nya untuk Jokowi. FBR berpindah dukung δšš |

The average of tweet length in our data is 127 character, even though current tweet limit character is 280. The maximum tweet length is 153 character and the minimum tweet length is 26 character. The tweet’s length distribution can be seen in figure 1.

![Tweet length distribution](image)

**Figure 1.** Emotion tweet’s length

4. **Research Method**

Generally, the process of this study divide into two big parts. First part is pre-processing data and second part is classification process. The pre-processing divide into extracting and selecting feature. The extracting feature process will use basic extracting feature which consist in removing unimportant or disturbing elements [16] such as URL, username, number, punctuation, ASCII character, and single letter. In addition, we will remove stopword and stemming. For stemming process we use sastrawi stemmer for stemming indonesian text. For this step, we will compare the result if we use stemming process and if we do not use it, because a study suggest that stemming does not give any significant result if the data collected from twitter [22]. The selecting feature, we use tf-idf method for weighted feature which already extracted. The second part of this study, we will perform classification process. We will compare two classifier, Naive Bayes Classifier and SVM, for classifying the emotion. This all processes can be seen in figure 2.

![Step-by-step text classification](image)

**Figure 2.** Step-by-step text classification
5. Result and Discussion

As suggested the steps defined on section 4, in pre-processing process we study the result either use stemming step or not. Each of the pre-processing result, we use two different classifier for classifying emotion. These classifiers are Naive Bayes and SVM (RBF kernel). We use cross-validation 10 folds for evaluating each model. The evaluation model of each classifier can be seen in table 2.

Table 2. Evaluation model classification

| Preprocessing  | Classifier    | Accuracy | Precision | Recall | F-measure |
|----------------|---------------|----------|-----------|--------|-----------|
| Without stemming | Naive Bayes (cv – 10 folds) | 46% (+/- 0.12) | 0.465 | 0.456 | 0.451 |
|                | SVM (cv–10 folds) | 46% (+/- 0.09) | 0.475 | 0.456 | 0.452 |
| With stemming  | Naive Bayes (cv – 10 folds) | 47% (+/- 0.07) | 0.487 | 0.471 | 0.468 |
|                | SVM (cv–10 folds) | 0.47% (+/- 0.09) | 0.485 | 0.471 | 0.467 |

From those evaluation model, we can see that Naive Bayes Classifier and SVM have no significant result. This support many studies which suggested Naive Bayes Classifier and SVM have good performance for text data.

As we can see stemming process is not giving any significant improvement in accuracy. This result same with the study as we mention before [22]. If the data collect from social media twitter, stemming process will not give any significant improvement in evaluation model. It does not give any significant result because indonesia tweet is not use formal writing and use slangs or abbreviation. This informal writing style can not be handled by stemming process.

Accuracy of each evaluation model is low, less than 50%. Those models have high missclassification class. Table 3 showed confusion matrix of one of the model, Naive Bayes Classifier with stemming.

Table 3. Confusion matrix - naive bayes classifier with stemming

| Predict | Anger | Sad | Joy | Surprise |
|---------|------|-----|-----|----------|
| Actual |      |     |     |          |
| Anger   | 53   | 20  | 12  | 15       |
| Sad     | 20   | 36  | 19  | 25       |
| Joy     | 13   | 17  | 58  | 10       |
| Surprise| 22   | 25  | 12  | 40       |

From table 3 confusion matrix, anger and joy class have high accuracy in prediction but sad class has lowest accuracy. Anger class has the most missclassification in sad class, sad class has the most missclassification in surprise class, joy class has the most missclassification in sad class, and surprise class has the most missclassification in sad class. So we can said that anger, joy, and surprise class has the most missclassification in sad class. It happen because features in sad class overlap with other three classes. 10 highest feature of each class can be seen in tabel 4. Anger, sad, joy, and surprise class have overlap feature such as ‘prabowo’, ‘jokowi’, ‘dukung’, and many others. These overlap features which have high score in weighted features can affect the developing model classification in classification process. The chance of missclassification error will be greater because the high rank features in each class does not unique feature. While there are many overlap features, anger and joy have more unique features than the other two classes. This help improve accuracy of its class.
Table 4. Ten highest feature of each class

| Rank | Anger  | Sad   | Joy   | Surprise |
|------|--------|-------|-------|----------|
| 1    | prabowo| jokowi| prabowo| prabowo  |
| 2    | jokowi | prabowo| jokowi| jokowi  |
| 3    | gak    | dukung| menang| dukung   |
| 4    | dukung | perintah| indonesia| presiden |
| 5    | fpi    | malu  | pilih | indonesia |
| 6    | kuasa  | rakyat| dengannmusikkitasatu| pilih |
| 7    | menang| indonesia| dukung| agum     |
| 8    | rakyat | negara| rakyat| jaman    |
| 9    | hoax   | percaya| musik| heboh    |
| 10   | orang  | ya    | sandi | gak      |

6. Conclusion
This study of emotion mining has shown classification method can predict emotion of tweet’s writer. This study has support that stemming step in text pre-processing of Indonesian tweets does not have significant improvement of evaluation model classification [22]. In classification process, Naive Bayes Classifier and SVM have not outperform one another. Overall model accuracy is not reach high accuracy because there are many misclassification. The evaluation model showed anger and surprise have high accuracy of prediction and sad has the lowest accuracy of prediction. One of the reason of this result is because of each class does not have many unique features. These unique feature will represent of its classes. In future study, we suggest similar study collect more data so each classes have rich unique features.

7. References
[1] suara.com, "Empat Hashtag Ini Rajai Debat Pilpres Semalam,” Suara.com, 31 March 2019. [Online]. Available: https://www.suara.com/tekno/2019/03/31/084338/empat-hashtag-ini-rajai-debat-pilpres-semalam. [Accessed 11 August 2019].
[2] A. Ceron, L. Curini, S. Iacus and G. Porro, "Every tweet counts? How sentiment analysis of social media can improve our knowledge of citizens’ political preferences with an application to Italy and France,” New media & society, vol. 16, no. 2, pp. 340-358, 2014.
[3] P. Burnap, R. Gibson, L. Sloan, R. Southern and M. Williams, "140 characters to victory?: Using Twitter to predict the UK 2015 General Election," Electoral Studies, vol. 41, pp. 230-233, 2016.
[4] U. Yaqub, S. Chun, V. Atluri and J. Vaidya, "Sentiment based analysis of tweets during the us presidential elections," in InProceedings of the 18th annual international conference on digital government research (pp. 1-10), ACM, 2017.
[5] M. Ebrahimi, A. H. Yazdavar and A. Sheth, "Challenges of sentiment analysis for dynamic events," IEEE Intelligent Systems, vol. 32, no. 5, pp. 70-75, 2017.
[6] A. Khatua, A. Khatua, K. Ghosh and N. Chaki, "Can# Twitter_trends predict election results? Evidence from 2014 Indian general election,” in 2015 48th Hawaii international conference on system sciences (pp. 1676-1685), IEEE, 2015.
[7] Y. Mejova, P. Srinivasan and B. Boynton, "Gop primary season on twitter: popular political sentiment in social media," in InProceedings of the sixth ACM international conference on Web search and data mining (pp. 517-526), ACM, 2013.
[8] T. Yasseri and J. Bright, "Can electoral popularity be predicted using socially generated big
data?," it-Information Technology, vol. 56, no. 5, pp. 246-253, 2014.

[9] P. Ekman, W. V. Friesen and P. Ellsworth, Emotion in the human face: Guidelines for research and an integration of findings, Pergamon Press, 1972.

[10] R. Plutchik and H. Kellerman, Emotion: theory, research and experience, New York: Academic press New York, 1986.

[11] A. Yadollahi, A. Shahraki and O. Zaiane, "Current state of text sentiment analysis from opinion to emotion mining," ACM Computing Surveys (CSUR), vol. 50, no. 2, p. 25, 2017.

[12] R. E. Jack, W. Sun, I. Delis, O. G. B. Garrod and P. G. Schyns, "Four not six: Revealing culturally common facial expressions of emotion.," Journal of Experimental Psychology: General , vol. 145, no. 6, p. 708, 2016.

[13] S. Dhawan, K. Singh and D. Sehrawat, "Emotion mining techniques in social networking sites," International Journal of Information & Computation Technology, vol. 4, no. 12, pp. 1145-53, 2014.

[14] S. Vijayarani, M. Ilamathi and M. Nithya, "Preprocessing techniques for text mining-an overview," International Journal of Computer Science & Communication Networks, vol. 5, no. 1, pp. 7-16, 2015.

[15] A. Krounska, C. Troussas and M. Virvou, "The effect of preprocessing techniques on Twitter sentiment analysis," in 2016 7th International Conference on Information, Intelligence, Systems & Applications (IISA) (pp. 1-5), IEEE, 2016.

[16] G. Angiani, L. Ferrari, T. Fontanini, P. Fornacciari, E. Iotti, F. Magliani and S. Manicardi, "A Comparison between Preprocessing Techniques for Sentiment Analysis in Twitter," KDWeb, 2016.

[17] L. Wikarsa and S. Thahir, "A text mining application of emotion classifications of Twitter's users using Naive Bayes method," in 2015 1st International Conference on Wireless and Telematics (ICWT) (pp. 1-6), IEEE, 2015.

[18] A. Susanti, T. Djatna and W. Kusuma, "Twitter's Sentiment Analysis on Gsm Services using Multinomial Naive Bayes," Telkomnika, vol. 15, no. 3, pp. 1354-1361, 2017.

[19] T. Fatyanosa and F. Bachtiar, "Classification method comparison on Indonesian social media sentiment analysis," in 2017 International Conference on Sustainable Information Engineering and Technology (SIET) (pp. 310-315), IEEE, 2017.

[20] I. Windasari, F. Uzzi and K. Satoto, "Sentiment analysis on Twitter posts: An analysis of positive or negative opinion on GoJek," in 2017 4th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE) (pp. 266-269), IEEE, 2017.

[21] N. Putranti and E. Winarko, "Analisis sentimen twitter untuk teks berbahasa Indonesia dengan maximum entropy dan support vector machine," IJCCS (Indonesian Journal of Computing and Cybernetics Systems), vol. 8, no. 1, pp. 91-100, 2014.

[22] A. Hidayatullah, C. Ratnasari and S. Wisnugroho, "Analysis of Stemming Influence on Indonesian Tweet Classification," Telkomnika, vol. 14, no. 2, pp. 665-673, 2016.