

A CLASSIFICATION MODEL TO IDENTIFY PUBLIC OPINION ON THE LOCKDOWN POLICY USING INDONESIAN TWEETS

¹ARMINDITYA FAJRI AKBAR, ²ARIS BUDI SANTOSO,
³PRABU KRESNA PUTRA, ⁴INDRA BUDI

Universitas Indonesia, Department of Information Technology, Jakarta, Indonesia

E-mail: ¹arminditya.fajri@ui.ac.id, ²aris.budi@ui.ac.id, ³prabu.kresna@ui.ac.id, ⁴indra@cs.ui.ac.id

ABSTRACT

The spread of the Pneumonia Coronavirus Disease 2019 (COVID-19) or Corona virus has affected several industrial sectors in Indonesia, particularly in the tourism and economy sector. Corona virus has been declared by the World Health Organization (WHO) as a pandemic that has spread to various parts of the world including Indonesia. In this regard, the Government of the Republic of Indonesia then declared the Corona virus as a non-natural national disaster. The Case Fatality Rate (CFR) of the Corona virus is 8.37%, placing Indonesia as one of the countries with the highest mortality ratio in the world. Currently, the Government of Indonesia has not implemented a lockdown policy, but there are some people who deplore the government's firmness in imposing the policy and there are also those who support the government for not making the lockdown status decision. Therefore, the lockdown is still a debate in the public. This can be read on social media Twitter, where many people express their opinions about the lockdown policy in Indonesia. Based on this polemic, this research has obtained a classification model that can differentiate between pro and contra tweets on the lockdown policy topics using Indonesian tweets. By using the Bernoulli NB algorithm as a classification model, an optimal value with the highest f-measure score of 88,57% was obtained. This model can be used to assess the effectiveness of communication in implementing lockdown policy to slow the spread of COVID-19 because it can identify public opinion about the trends in supporting or rejecting the lockdown policy.

Keywords: COVID-19, Lockdown, Sentiment Analysis, Supervised Learning, Text Classification

1. INTRODUCTION

Pneumonia Coronavirus Disease 2019 (COVID-19) or Corona virus is a disease of lung inflammation caused by Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) which can cause respiratory system disorders, acute pneumonia, even death [1]. The spread of the Corona virus has affected several industrial sectors in Indonesia, particularly in the tourism and economy sector. The Ministry of Tourism and Creative Economy, Wishnutama Kusubandio, said that Indonesia's tourism sector has a potential to lose up to US\$ 4 billion (Rp. 57 trillion) with the closure of several tourist attractions [2]. In addition, Perry Warjiyo as a Governor of Bank Indonesia said the economic growth is projected to decline by 3% from 2020 estimates due to global supply chain disruption that is caused by access closure in several countries [3].

Corona virus has been declared by the World Health Organization (WHO) as a pandemic that has spread to various parts of the world including

Indonesia [4]. The Government of the Republic of Indonesia later declared the Corona virus as a non-natural national disaster [5]. Based on Figure 1 that is sourced from Katadata, the spread of the Corona virus began on March 02, 2020 with the identification of 2 Indonesian citizens who were positively affected by the pandemic, as of March 18, 2020, there were 227 patients with 197 people were being treated, 19 people died, and 11 people were declared cured. The high Case Fatality Rate (CFR) of the Corona virus, which is 8.37%, putting Indonesia as one of the countries that has the highest mortality ratio in the world [6].

A lockdown can be one of the solutions as an effort to minimize the spread of Corona virus. A lockdown is a situation where people are not allowed to enter or leave an area due to an emergency [7]. There are several countries that have succeeded in applying lockdown, thereby reducing the number of COVID-19 infections, namely New Zealand, Australia, and Malaysia, with an average recovery rate of the three countries reaching 89.6% [8].

At the moment, the Government of Indonesia has not implemented a lockdown policy, but there are some people who deplore the government's

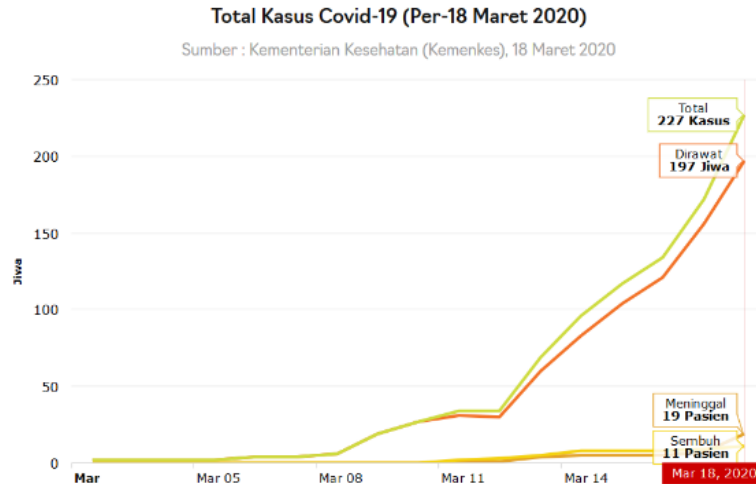


Figure 1: Total Cases of COVID-19 in Indonesia

firmness in imposing the policy and there are also those who support the government for not making the lockdown status decision. Therefore, the lockdown is still a debate in the public. This can be read on social media Twitter, where many people express their opinions about the lockdown policy in Indonesia.

Based on this polemic, this research was conducted with the aim of creating a classification model that could identify opinions about the tendency of the public in supporting or refusing the lockdown policy by the Government of the Republic of Indonesia as a general description of their opinions in responding the handling of the Corona virus spread.

This paper consists of 5 sections and is organized as follows: Section 2 describes literature review in sentiment analysis studies. Our methodology is explained in section 3. We discuss experimental results in section 4 and finally, section 5 contains conclusions about this research.

2. LITERATURE REVIEW

In this section, we limit ourselves to government subjects that use sentiment analysis since it represents the main focus of this research. E. D'Andrea et al [9] research presents an approach to monitor Italian public opinion from tweet's analysis, by referring to the topic of vaccination. The topic became controversial, due to the alleged relationship between autism and vaccines. This system was built to monitor negative public opinion about vaccines as a material for a quick decision

making. The approach proposed in this research consists of three phases, which are a collection of tweets, text elaboration, and opinion mining. In addition, the module development used the Java programming language and Weka (Waikato Environment for Knowledge Analysis). The algorithm used in this research is Simple Logistic Classifier with the resulting level of accuracy of 75.5%. We followed a similar approach to opinion mining in our study. However, the detailed steps in the experiment are different in our case (i.e. public opinion on the lockdown policy).

M. N. Aziz et al [10] research focuses on method developing to detect the government service satisfaction in the city of Surabaya on Identity Cards (KTP) using big data analysis on Twitter. The approach is carried out in two ways, namely the unsupervised classification and supervised classification approach to find the best method. The tweet data that will be used is processed using the preprocessing approach first. In addition, the Latent Dirichlet Allocation (LDA) approach is used to look at the topics that tend to be strong which influence negative or positive opinions. The results of this research found that the approach with supervised classification gives more satisfying results with a comparison of 75% and 53% accuracy for the unsupervised classification approach. The use of feature extraction using TF-IDF classification and Support Vector Machine (SVM) is able to build models that are better than the lexicon-based approach. Based on the results of this paper, we followed the approach using supervised

classification because of the high accuracy. However, the feature engineering process in this study uses a different approach, which uses a bag of words with binary values to build classification model.

N. A. Salsabila et al [11] is a research that focuses on compiling a list of terms and languages used daily in Indonesian. The results from compiling the glossary are then normalized into words that are more meaningful and are supplemented with information. The function of the lexicon dictionary is as a dictionary for text normalization steps and as a dataset to build text normalization models. The cumulative language or also known as "alay language" was collected as many as 3,592 words, coming from the Instagram comments column. The results of the evaluation of this research are the need for more extensive research in using the lexicon dictionary for other cases such as sentiment analysis, topic modeling and others in order to improve performance in Indonesian language social media analysis. The research is used as a reference in normalizing non-standard words into standard words because in Indonesian there is a term or colloquially called "alay language". Continuing the research, our study applies additional process of text normalization to correct word writing errors in a sentence so that preprocessing result may be more optimal. In addition, this study experimented with the Naïve Bayes Classifier, Decision Tree, Bernoulli NB, and Complement NB algorithms to build a classification model using bag of words with binary values. These algorithms have not been used in previous studies.

3. METHODOLOGY

The purposed workflow of making the sentiment analysis models in this research is shown in Figure 2. This research begins with acquiring text, preprocessing data, feature engineering, classification model development, model evaluation, and data visualization.

3.1 Acquiring Text

In this research, the dataset used is a tweet that contains opinions about people's tendencies in supporting or refusing lockdown policies as exemplified in Table 1. Crawling data is conducted on social media Twitter using the Python programming language by utilizing the Twitter API Developer services.

The tweet data collection period began on March 14, 2020 until March 27, 2020 with 2,321 tweets of data successfully collected. The query used to

retrieve the tweet data is "*pemerintah AND (covid OR covid-19 OR covid19 OR corona OR korona) AND (lockdown OR 'lock down') -filter:links -filter:images -filter:native_video -filter:retweets*". In English, query will be looked like "*government AND (covid OR covid-19 OR covid19 OR corona) AND (lockdown OR 'lock down') -filter:links -filter:images -filter:native_video -filter:retweets*".

Then, all the crawling data in JSON format is saved in *Comma-Separated Values* (CSV) format by only utilizing the text content of the tweet. Every tweet in the CSV file will be labeled in the "pro" or "cons" category.

Table 1: Dataset Sample that Expresses the Public Opinion

Tweet	Classification
RasaUtang: Aku mah salah satu orang yang menolak #lockdown untuk Indonesia (<i>I'm one of the people who refused #lockdown for Indonesia</i>). Bagaimana bisa menutup segala aktivitas publik, bandara, pemerintah, sekolah dan segala macamnya hanya untuk menghadapi virus Corona. Akan menciptakan panik sosial yang sangat berlebihan.	Cons
glnbgskara: @queensky021 Gw lebih setuju lockdown total, pemerintah nyatanya kecolongan kok soal virus ini (<i>I agree to a total lockdown, the government has actually been cheated about the spread of this virus</i>), masih ingat ketika peneliti havard diragukan soal penyebaran virus Corona di Indonesia yang ditepis kemenkes, efek domino lockdown harus diterima mau gamau yang penting isolasi total dulu	Pro
dioryaldino: Hikmah dari corona adalah kantor gue flexitime seperti impian gue. Haha. Gue menolak lock down karena jangan sampe ekonomi lumpuh total (<i>I refused lock down to avoid total economic paralysis</i>). Kasian yg penghasilan harian. Pemerintah gak akan sanggup cover itu semua. WFH, Pembatasan sosial dan tetap bersih masih opsi terbaik saat ini.	Cons
nirda_: aku setuju kalo Indonesia segera membuat kebijakan lockdown (<i>I agree that Indonesia immediately makes a lockdown policy</i>), Pemerintah tolong dong lebih baik mencegah sejak dini daripada semakin parah nantinya 😊 #LockDownIndonesia #COVID19	Pro

3.2 Preprocessing Data

The dataset that has been obtained is then carried out preprocessing to avoid the opinion data that is not

ready to be processed as there is interference (noise) [12]. The stages in this preprocessing are as follows:

3.2.1 Case Folding

All the letters in the text are converted into the lowercase format. This is conducted because Python is a programming language that is case sensitive, and the dictionary of words used is in all lowercase text format [13].

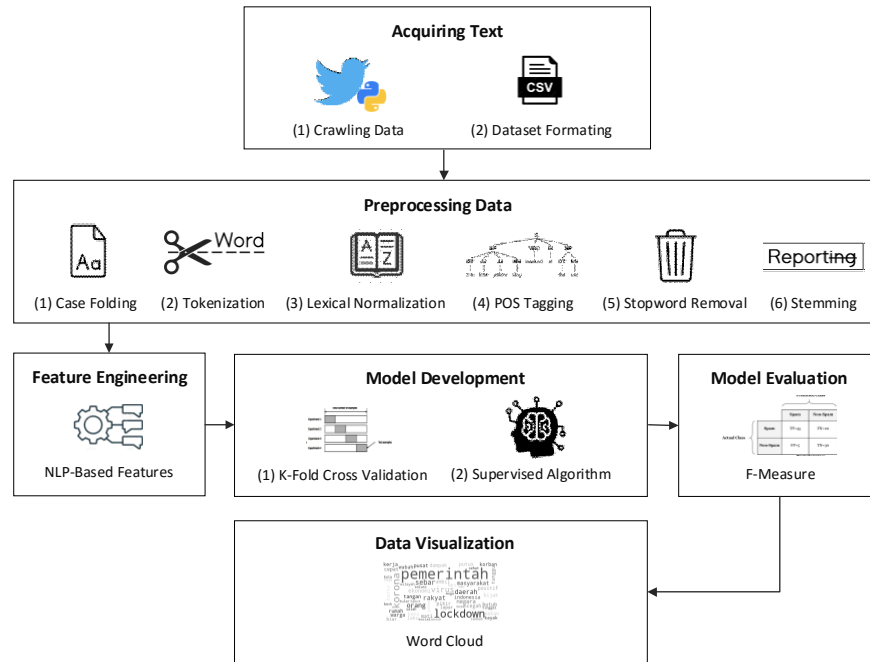


Figure 2: Proposed Workflow

3.2.2 Tokenization

A set of characters in an Indonesian text that is broken down into units of words (unigram) using a regular expression to assist in the next preprocessing step.

3.2.3 Lexical Normalization

The normalization process is carried out to convert nonstandard words into canonical forms so that the process of stopword removal and stemming produces a maximum output [12].

3.2.4 POS Tagging

Specifying the part-of-speech of a sentence using Flair NLP to determine nouns, verbs, adjectives, and others.

3.2.5 Stopword Removal

Erasing words that have no meaning or have a low frequency in general. In addition, eliminating punctuation, retweet, hashtag, and mention in tweets to reduce the number of words stored in the token list.

3.2.6 Stemming

Change all the words in the document into basic words by removing all the affix using the Literature library.

3.3 Feature Engineering

Feature engineering is compiling flat feature from words or terms produced in the preprocessing data process using NLTK. The types of words used in flat feature are nouns, verbs, and adjectives. The choice of words is conducted to eliminate the noise feature so that the runtime model that will be generated more quickly [12]. Afterwards, arranging the distribution of terms on the flat feature based on the frequency of occurrence and taking the first 3,000 of the terms distribution to make a certain classifier feasible.

Figure 3 illustrates the anatomy of the feature set that will be used in the classification modelling as a training set.

```

1  [
2  (
3  {
4  'pemerintah': True,
5  'lockdown': True,
6  'korona': False,
7  'virus': False,
8  'sebar': False
9  },
10 'kontra'
11 )
12 ]
    
```

Figure 3: Anatomy of the Feature Set

Feature set is formed as list which contains sets of binary terms (lines 3-9) and labels of tweets (line 10).

3.4 Model Development

Text classification can be done by training the classifier model first using the training set owned. A training set is formed by looking for features in every tweet. Then, a feature set is used to form trend of "pro" or "cons" sentiments. Some steps that need to be done next are as follows:

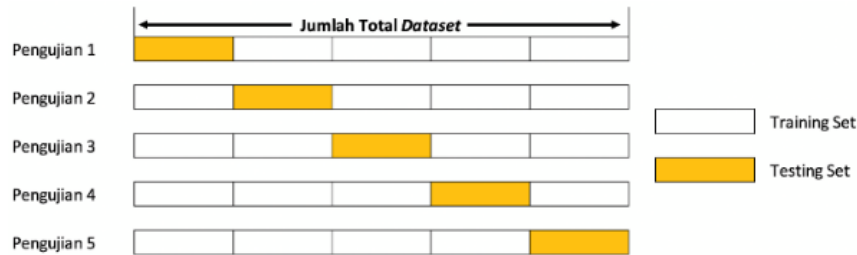


Figure 4: Iteration Process with 5-Fold Validation

3.4.1 K-fold cross validation training

Dividing the training set to train the classification model and the testing set to test the classification model. As explained in Figure 4, the position of the slices of the testing set will change according to the specified number of folds (k). In this research, the specified k value is 5-fold (five folds), which means that there will be 5 times of the classification model testing.

3.4.2 Supervised machine learning model

The final step is to build a model where the machine learning model is trained using the allocated training set. Algorithm variation from supervised machine learning (Naïve Bayes Classifier, Decision Tree, Bernoulli NB, and Complement NB) will be used to compare the evaluation value based on the testing using a testing set.

3.5 Model Evaluation

After the text classification process using several supervised machine learning models is complete, the next process is to calculate the accuracy of the classification model using f-measure. F-measure is an assessment that combines the results of precision and recall in a situation where the dataset used has an imbalanced proportion. To determine the weight of the classification model, a calculation is performed using the following formula:

$$F - \text{Measure} (F) = \frac{2TP}{2TP + FP + FN} \quad (1)$$

4. EXPERIMENTAL RESULT

This research uses a dataset obtained from Twitter totaling 915 relevant data from a total of 2,321 tweets. The relevant data used in this research is the result of manually filtering the tweet data. Relevant Tweets contain public opinion that only relate to the lockdown policy with examples listed in Table 2.

Table 2: Example of Relevant and Irrelevant Tweets

Subject	Tweet
Relevant Tweet	@Nutrisi6 @Dennysiregar7 Info yg transparan saja ke publik dari pemerintah, mana daerah yg terparah terparah covid-19 , itu yg di lockdown dan konsentrasi penuh utk penanggulangannya, sehingga seluruh rakyat bisa mempersiapkan mental dan menjaga kesehatan, kebersihan diri @Nutrisi6 @Dennysiregar7 Just transparent information to the public from the government, where the worst affected areas are covid-19, then the area is locked down and in full concentration to overcome it, so that all people can prepare mentally and maintain personal health & hygiene
Irrelevant Tweet	Jika pemerintah membuat kebijakan lockdown untuk menekan persebaran corona berarti aku juga harus buat kebijakan lock your heart to me biar nanti kalo ada tikungan dari si dia kamu dengan tegas menolaknya ☐ If the government makes a lockdown policy to suppress the distribution of corona, it means that I also have to make a policy of locking your heart to me so that if there is a turn from her then you can firmly reject it ☐

Furthermore, the labeling process in the dataset is conducted by a simple voting method from three annotators to analyze whether the sentence of the tweet has a “pro” or “cons” opinion.

From all the 915 tweets, there were 555 pro tweets and 360 contra tweets of the lockdown policy so that it can be concluded that the data taken from Twitter illustrates the people who tend to support the implementation of the lockdown. The initial distribution of the tweet data will be used as a test and training data in the development model process. The proportion of data used in this research is represented in Figure 5.

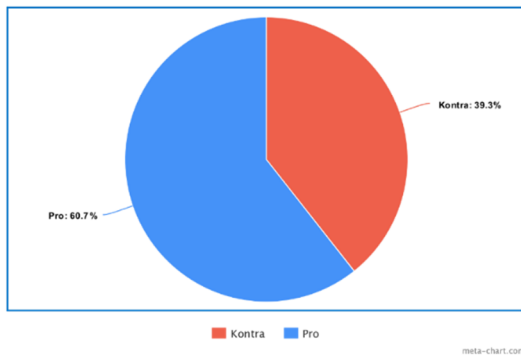


Figure 5: Iteration Process with 5-Folds Validation

The dataset used in this research are tweets and sentiments. Before analyzing the sentiments that have been collected, the step that needs to be done is preprocessing data that aims to reduce the dimensions of words that have no effect on the results of the data processing so that the system can process the data classification more quickly with more accurate results [12]. The preprocessing data section consists of six stages with the results shown in Table 3.

Table 3: Normalization Result From Tweet Samples

Tweet Sample	Normalization Result
@innovesian Kalo gak lockdown, sebulan lagi uang mrka akan habis untuk biaya pemakaman sanak saudara mereka yg meninggal karna corona. Lockdown cukup sebulan aja, cicilan telat karna wabah gini	kalaupun tidak lockdown sebulan lagi uang mereka akan habis untuk biaya pemakaman sanak saudara mereka yang meninggal karena corona lockdown cukup sebulan saja cicilan telat karena wabah begini masih bisa

Tweet Sample	Normalization Result
masih bisa diurus @ojkindonesia Makan sebulan rakyat kecil hrsnya bisa ditanggung pemerintah	diurus makan sebulan rakyat kecil harusnya bisa ditanggung pemerintah

After preprocessing data and compiling the feature set containing 2,231 binary term frequency with labels ("pro" or "cons") then a validation (model development) is carried out with k-fold cross validation using 5-fold where a cross validation process is done with 5 times iteration. Each iteration or fold contains of 732 datasets used as a training data and 183 datasets used as a test data that is classified with supervised algorithms, namely Naïve Bayes Classifier, Decision Tree, Bernoulli NB, and Complement NB.

Table 4: Cross Validation Results for Model Evaluation

Iteration						Avg.
	1	2	3	4	5	
Naive Bayes Classifier	71,64%	78,26%	72,00%	81,06%	87,22%	78,036%
Decision Tree	71,09%	71,68%	70,42%	80,00%	79,07%	74,452%
Bernoulli NB	71,37%	76,60%	74,89%	83,40%	88,57%	78,966%
Complement NB	71,79%	79,00%	73,20%	80,91%	86,69%	78,318%

Based on Table 4, the optimal results are obtained from the classification model using the Bernoulli NB algorithm in the fifth iteration with a score of 88,57%. The results of the model evaluation are obtained using the f-measure formula. From this research, the evaluation value of the model is influenced by the number of tweets because there will be more features to be trained if there are more tweets on the machine learning algorithm [14].

To prove that the resulting model can analyze sentiments well, a model testing is performed. Model testing process is carried out by the inclusion of new tweets from Twitter related to public opinion on the lockdown policy for an automatic sentiment analysis. The test results can be seen in Table 5.

Table 5: Test Result of the Classification Model

No	Tweet	Actual	Predicted
1	kemarin emang sy setuju sama lockdown. kirain bantuannya merata, eh nggak. jadi saya lebih baik menentang aja. (<i>yesterday I agreed with lockdown. I thought social assistance was evenly distributed, apparently not. so I would rather just oppose that.</i>) toh pemerintah jg gak peduli. yang penting sy tetep menjalankan protokol kesehatan.	Cons	Pro
2	Setuju untuk gak lockdown dulu. Pemerintah harus ngelakuin transparansi informasi dulu. (<i>Agree to not lockdown first. The government must first transparency the information.</i>) Maksimalin. Kalo penyebarannya semakin parah baru lockdown. Contoh pengecekan covid-19 dulu digratisin. Rajin2 sebar informasi, biar yang tolol2 gak seenaknya nyebarin hoax.	Cons	Cons
3	Salah satu solusi yang harusnya diambil pemerintah itu harusnya menolak kunjungan wna dari manapun buat masuk ke indonesia & kembalikan mereka segera ke negara asalnya. Trus lanjut dah pemerintah lockdown, tangani & cegah covid 19 di masyarakat. (<i>Then the government should implement lockdown policy, handle and prevent covid 19 in the public.</i>)	Pro	Pro
4	rakyat ga bs disiplin kesalahan sepenuhnya ada pd rakyat indonesia, pemerintah yg tdk mendukung & melaksanakan lockdown total dari awal pandemik covid 19 terdeteksi (<i>the fault is entirely with the people of Indonesia because they cannot be disciplined, governments that do not support & carry out a total lockdown from the start of the covid 19 pandemic are detected</i>). klu dari awal lockdown, ramadhan akan aman & .nyaman. (<i>if from the beginning do lockdown, ramadhan will be safe & .comfortable.</i>)	Pro	Cons
5	Jangan sampailah di Lockdown khususnya di Kota Jakarta, Kota itukan Central Business Indonesia, masa Pusat Perekonomian Indonesia di Lockdown.. (<i>Do not let Lockdown apply especially in Jakarta, the city is Central Business Indonesia, how come the Indonesian Economy Center is applied Lockdown..</i>) Jadi marilah teman-teman kita patuhi peraturan yg pemerintah sudah berikan dan marilah kita saling mendukung utk melawan Virus Corona	Cons	Cons

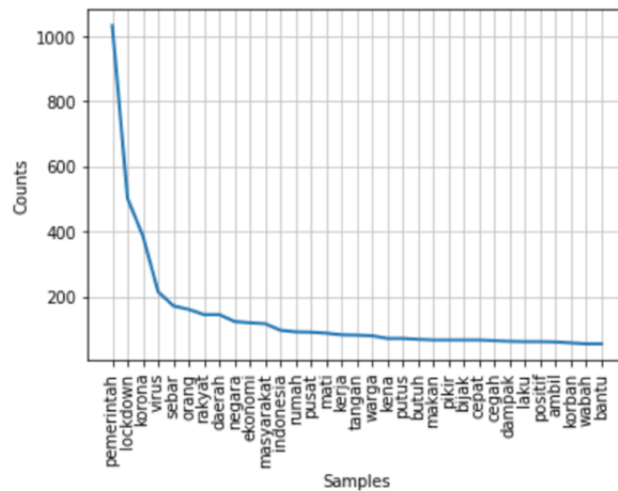


Figure 6: Frequency Information of the Appearance of Each Token

Based on Table 5, it can be proven that the classification model is quite good in analyzing sentiments from random tweets that do not have labels. From the 5 new tweets, there are 2 tweets that do not match the prediction. Thus, it can be concluded that the classification model works pretty well.

Figure 6 illustrates the frequency distribution of the appearance of each tokens or words in the dataset used in this research. Figure 6 shows that the terms

pemerintah (government), *lockdown*, *korona* (corona), *virus*, and *sebar* (spread) are some of the most common terms in the public posts related to the *lockdown* policy on Twitter. The presence of terms related to *mati* (die), *kerja* (work), *kena* (hit), *putus* (break), *butuh* (need), *makan* (eat), *positif* (positive), and *korban* (victim) illustrates anxiety along with the rapid spread of the Corona virus. The appearance of words such as *pikir* (think), *bijak* (wise), *cepat* (fast), *cegah* (prevent), *ambil* (take), and *bantu* (help) is an indication of the desire to act immediately.

- [10] M. N. Aziz, A. Firmanto, A. M. Fajrin and R. V. H. Ginardi, "Sentiment Analysis and Topic Modelling for Identification of Government Service Satisfaction," *International Conference on Information Technology, Computer and Electrical Enginerring (ICITACEE)*, vol. V, pp. 125-130, 2018.
- [11] N. A. Salsabila, Y. A. Winatmoko, A. A. Septiandri and A. Jamal, "Colloquial Indonesian Lexicon," *International Conference on Asian Language Processing (IALP)*, pp. 226-229, 2018.
- [12] S. Gupta and A. Gupta, "Dealing with Noise Problem in Machine Learning Data-sets: A Systematic Review," *ELSEVIER*, vol. 5, pp. 466-474, 2019.
- [13] Python Software Foundation, "PEP 235 - Import on Case-Insensitive Platforms," [Online]. Available: <https://www.python.org/dev/peps/pep-0235/>. [Accessed 05 Juni 2020].
- [14] Medium, "Rants on Machine Learning," 7 Ways to Improve your Predictive Models, [Online]. Available: <https://medium.com/rants-on-machine-learning/7-ways-to-improve-your-predictive-models-753705eba3d6>. [Accessed 05 Juni 2020].