Sentiment Analysis of Face-to-face Learning during Covid-19 Pandemic using Twitter Data

Ghanim Kanugrahan[1] and Alfan Farizki Wicaksono[2]
Faculty of Computer Science
Universitas Indonesia
Jakarta, Indonesia
Email: [1]ghanim.kanugrahan@ui.ac.id, [2]alfan@cs.ui.ac.id

***Abstract*— Covid-19 pandemic has massive impacts on the activity of human in the world, including in Indonesia. To reduce the transmission of the virus, Indonesian government issues a policy to restrict daily public activities, affecting key national sectors, such as education systems. All learning activities are switched from the conventional face-to-face mode to being remote via the use of the Internet. After the pandemic begins to subside, the government then plans to reopen all schools and to allow face-to-face learning. However, this decision has sparked controversy in the social media, including Twitter. This paper describes a methodology to perform sentiment analysis on a collection of tweets that are in connection with the restart of the face-to-face learning mode. In particular, our experiments using hand-crafted features based on the tweets demonstrate that data-driven models are useful for automatic sentiment orientation classification on Twitter data. The best model achieved in this study has 69,1% accuracy, 68.6% precision, 69.1% recall, and 67,8% F1-Score. This result is achieved by using unigram, Support Vector Machine, and tweet + number of words (count) feature combinations.**

Keywords*— *machine learning, SVM, ANN, Covid-19, face-to-face learning

I. Introduction

At the end of 2019, China reported to the World Health Organization (WHO) that a virus had begun to spread. A few months later, the virus quickly spreading and make the world shocked by how quickly the widespread of the virus. The virus later named as Covid-19 or novel coronavirus [1]. By March 2020, Covid-19 already spread into various parts of Indonesia. Carried by air, it made the virus easy to infect others. This phenomenon has caused the government to act quickly to break the chain of Covid-19. For this reason, the Indonesian government has taken the decision to limit all forms of activities related to people gathering and crowds, including annual holiday exodus [2]. This policy resulted in the inhibition of various activities of Indonesian people, ranging from the economic sector to the field of education.

To overcome this challenge, Kementerian Pendidikan dan Kebudayaan (the Ministry of Education and Culture) launched a distance learning program [3][4]. With that program, all educational institutions began to switch to distance learning by utilizing Internet facilities. Even so, the program adopted by the Indonesian government is considered to be ineffective [5]. This problem is due to various obstacles experienced by teachers and students. These obstacles include the lack of experience and facilities to carryout distance learning.

With these conditions, the Indonesian government began to accelerate preparations for the return of face-to-face learning. This process begins by conducting trials of face-to-face learning in various regions [6]. From the trials, the government hoped they can determine the best policy for the re-implementation of face-to-face learning to the fullest. However, this policy raises various public response. Many of them concerns about the facilities and protocols [7]. Other sided with the Minister that face-to-face learning is important for the students [8]. From this problem, Indonesian people's opinion is divided into two major groups. The government needs to address the problem and find the best solution. One of them is by observing people's opinion about the matter. With the debate about face-to-face learning spread into social media platforms, we can use the social media as the data source as sample to observe the said problem.

One of the social media platforms that popular amongst the Indonesian is Twitter. By using Twitter, its user can express their thought through short post named tweet. By using data from users' tweet, we can build a model to predict users' sentiment about face-to-face learning by using machine learning approach. Basically, it is done by classifying tweets into several class, which are pro or positive, contra or negative, and neutral labels. In this study, we will use Machine Learning approach to identify Twitter users' sentiment on the matter. We will also evaluate which data features that increase the model quality. By using machine learning approach, we can build a model that automatically classify new tweets based on the acquired tweets.

There are two research questions that can be inferred from this research which can be seen as follows.

RQ1: To what extent the sentiment of face-to-face learning during Covid-19 pandemic can be predicted using machine learning based on hand-crafted features?

RQ2: In the context of RQ1, what kind of features will be useful for the sentiment prediction task?

II. Background

A. Sentiment Analysis

Sentiment Analysis is a study that focuses on analyzing people's emotions, opinions, and responses from a text [9]. It is also known as opinion mining and often used to analyze a public response to an event or phenomenon that has occurred or will occur. Sentiment analysis is included in Natural Language Processing category, which is a study that intended to analyze and understand the hidden pattern that contained in a text data. With that approach, researchers can take the essence of a text automatically with little human intervention.

Generally, sentiment analysis is carried out by an organization to obtain feedback and criticism of a topic, where opinions that develop in the community will be divided into positive and negative opinions [10]. Then, the data can be processed with qualitative and quantitative research. In quantitative research, statistical calculations will be carried out to find new information hidden within the collected data. With this new information, it is hoped that new insights can the organization to make policy-making decisions.

B. Support Vector Machine

Support Vector Machine, or SVM for short, is a method of a supervised machine learning that use to predict some class by using the data pattern [11]. With the data pattern, it will look for the best way to separate the class of the data. In order to find the best class separation, it needs to find the support vector of each class. Support vector is the farthest data from each class. Then, we will draw the line between the support vector from each data to separate each class. This separation will generalize the value for each class. The line that separating each class is called a Hyperplane [12].

C. Artificial Neural Network

Artificial Neural Network, or ANN for short, is inspired by the usage of human brain [13]. It is represented by some nodes and links that connected each nodes to other nodes. Artificial Neural Network process the data, then train itself with the acquired data. From that process, Artificial Neural Networks can predict the result from the data processed by finding the similarity from the data. Artificial Neural Network consist of three layers. Each layer consists of nodes that contains information. These layers named input layer, hidden layer, and output layer [14].

The input layer will take the external data to be processed by the algorithm. Then, the result will be predicted in the output layer. Between the input layer and output layer, there is a layer named hidden layer. Hidden layer contains some value named bias. With bias, hidden layer will process the calculation needed for the process. There may be several hidden layers. With more hidden layer, it will increase the calculation needed for the result. Each layer is connected to other layers by numbers of channel from each nodes. These channels have some numerical value named weight. With the calculation of weight and bias, we can calculate the prediction using Artificial Neural Network. One of the most popular examples of ANN is Multilayer Perceptron (MLP). MLP is one of the simplest ANN applications that can be used in machine learning process [15].

D. Previous Study

Sentiment analysis can be done in several ways and algorithms. Those algorithms can be divided into two major parts, through unsupervised or supervised learning. There are several approaches that can be used on this research which were obtained from different literatures and journals.

In 2017, Mowery et al [16] conducted a study about the usage of sentiment analysis to identify student depression. They use Twitter data in form of user's tweets to be processed by using machine learning approach, which is SVM. In that study, they also compare the most contributing features to acquire highest accuracy. The results are they can use simple lexical features and reduced feature sets can produce comparable result to much larger datasets. However, the study did not specify on how the data was labelled. Compared to their study, our study focused on the simple addition of new features based on the new features.

In 2020, Sukma et al [17] conducted a study about the usage of sentiment analysis to identify people's reaction to the ratification of Omnibus Law in Indonesia, which became a hot topic in Indonesia at the time. They use Twitter data in form of user's tweets to be processed by using machine learning approaches. The algorithm used in their study are SVM, Naïve Bayes, and Decision Tree. In the study, they compare each of the algorithm's accuracy. The result is SVM have the highest F1-Score with 92.00%. However, they mentioned about examples of different topics in their study. In the same time, they also mentioned examples of positive and negative sentiments in their study. Another difference with their study is that they only used tweets to create the model in their research with no additional features in their dataset.

In 2017, Ozturk et al [18] conducted a study about the usage of sentiment analysis in education system. This method is used to review the performance of an education system in Anadolu University, named Anadolu University open and distance education system. The system has more than two million users, then their opinion about the system will steer on how the system will be developed.

By using Twitter API, the tweet within two weeks timespan is collected. Then, the data is processed using Language Detection API to process and filtering the tweets from many countries. Using Naïve Bayer Classifier (NBC) method, the tweets then classified and resulted in more negatives than neutral and positives. Finally, the words then visualized using Word Clouds method. This research barely explains how the data is processed using NBC method. Then, there is also lacks of explanation using the word cloud method. In the end, the accuracy using the proposed method has not reach 70%. Therefore, our study proposed a new method by using different machine learning algorithm with additional features to create a new model.

Based on those previous studies, we conclude that Twitter can be used as a data source to find out public sentiment on several topics. This process of sentiment analysis can be done using several approaches. One of them is by using machine learning and feature selection. In this study, we propose a new method by adding additional features. These features are extracted from the tweet.

III. METHODOLOGY

This study implemented few research stages that need to be performed for sentiment analysis. These stages are carried out in order to achieve the result that answers the following research questions as well as the goals or purposes to achieve. There are 4 steps for this study can be described as follows.

1. Problem Formulation

The first stages in this research is to do the formulation of the problems by investigate popular event pas several week. Then, we try to find out the problems that were faced is needed to be implemented at the beginning of this research. Besides, the methods that will be used to overcome those problems also needs to be done.

In this study, we acknowledge that after more than a year of Covid-19, Indonesian Government wanted to implement face-to-face learning program as usual. Many of them will express their opinion on social media platform, one of them is Twitter. From that phenomenon, we wanted to build a model to classify user's opinion using machine learning approach. Therefore, the purpose of this study is stated in the research questions, namely in Q1 and Q2.

2. Literature Study

After formulating the problems in the form of research questions, some exploration in terms of tools and theory about sentiment analysis is conducted for implementing the next stage of this research, which is literature study. This study mainly focused on the topic of sentiment analysis based on previous studies of Mowery et al [16], Sukma et al [17], Ozturk et al [18], and Shahnawaz et al [19]. Based on those literature, we formulate the best algorithm for this study.

3. Scenario Plan

In scenario plan, we built a framework based on the literature study that have been conducted. Firstly, we will gather the tweets from Twitter using Tweepy. The language is set to 'Indonesian' with several keywords like:

- Sekolah tatap muka (face-to-face school)
- Sekolah offline (offline school)
- Kuliah offline (offline class)
- Kuliah tatap muka (face-to-face class)

The features that will be collected are "tweet" and "date". Then, the data will be labelled by 3 different annotators. Their qualifications can be seen in Table I:

TABLE I. ANNOTATORS OF TWEETS

	Annotator 1	Annotator 2	Annotator 3
Age	23	24	24
Education	Bachelor degree	Bachelor degree	Bachelor degree
Major	Nursery	Computer Science	Computer Science

The annotators will give the tweets into one of the three labels. There are "-1" as "negatives", "0" as "neutral" and "1" as "positives". If there are different opinions about label of some tweets, the annotators will discuss until the labels are agreed. After that, the data is pre-processed using several pre-processing techniques, which are:

- Lower-Case

  Change all words into lowercase. E.g., "WoRlD" will be "world", turning 'W', 'R', 'D' into lowercase.

- Remove Mention

  Remove all mention and words started with '@'. E.g., "@fbi open up" will be "open up", removing '@fbi'.

- Remove Punctuation

  Remove all punctuation in the text. E.g., "Hola!!!@!" will be "Hola", removing '!' and '@'.

- Remove Elongated Words

  Remove repeated letter more than 2. E.g., "Maaaaaaaf bangeeet" will be "Maaf banget", removing repeated 'a' and 'e' making it no more than 2.

- Stopwords Removal

  Removing Indonesian stopwords. E.g., "dan terjadi lagi" will be 'terjadi', removing 'dan' and 'lagi'. This process uses Sastrawi library, which process Indonesian words based on the dictionary. This library is used to simplify the research process.

- Normalization

  Turn Indonesian words into basic forms. E.g., "kusimpan dihati" will be "simpan hati", turned 'kusimpan' and 'dihati' into its basic forms. This process also uses Sastrawi library.

- Tokenization

  Turn words into token.

The example of these process will be resulted in Table II

TABLE II. AFTER PRE-PROCESSING

tweet	date
aku senang	2021-04-08 22:54:21
tidaakk mauu	2021-04-08 08:54:21

After pre-processing method, we will add more features into the data. By adding new features to the data, it is expected to have better results [20]. After looking at the acquired data, it is concluded that the features that can be added are:

- Count

  This feature shows how many words in a tweet. The data is taken from "tweet" feature. It is expected that the same labels will have similar tweet length

- Hour

  This feature shows in which time it was tweeted. The data is taken from the hour value in "date" feature. It will be divided into 3 main times, "morning" (00:00:00-08:59:59), "afternoon" (09:00:00-16:59:59), and "evening" (17:00:00-23:59:59). It is expected that the same labels will have similar time when the tweets were posted.

After we added more features, we need to process each feature with different feature extraction process. First of all, we will apply One Hot Encoding to "hour" feature. This process is used to process categorical data from the dataset. The result can be visualized in Table III.

TABLE III. AFTER ADDED FEATURES

tweet	count	hour morning	hour afternoon	hour evening	Label
aku senang	3	0	0	1	1
tidaakk mauu	2	1	0	0	-1

After we added more features, the table will consist of "tweet", "count", "hour pagi", "hour siang", "hour malam", and "label". Then, we experimented on many feature combinations to discover which combinations will produce the best model. The experiments in this study are:

- Only tweet
- Tweet + count
- Tweet + hours
- Tweet + count + hours

After that, the data will be divided into training data and testing data. The proportion of the training data and testing data will be 80:20. Then its features will be extracted using TF-IDF with n-gram unigram and bigram. For the "count" feature, we will use StandardScaler to simplify the number in the dataset. After that, the dataset will be processed using SVM linear and ANN. the result evaluated using accuracy, precision, recall, and F1-Score (F1) metrics.

4. Model Evaluation

The last stage in this research is to analyze the results based on the confusion matrix that consists of accuracy (1), precision (2), recall (3), and F1 (4). To calculate those scores, we will use cross validation with ShuffleSplit, which divide the data into 5 parts by random. Then, these 5 data parts will be tested into each other. Finally, the data will be calculated using formulas as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

$$F1 = 2 \times \frac{Precision\ x\ Recall}{Precision + Recall} \tag{4}$$

From the formula (1), (2), (3), and (4), there are some abbreviations, such as TP, TN, FP, and FN. TP stands for True Positive, which means how many times the model can correctly predict the *true* label of a tweet. TN stands for True Negative, which means how many times the model can correctly predict the *false* label of a tweet. FP stands for False Positive, which means how many times the model falsely predicted the *true* label of a tweet, Finally, FN stands for False Negatives, which means how many times the model falsely predicted the *false* label of a tweet.

IV. Results and Discussion

A. Data Gathering

The data is collected from 10th of May 2021 until 8th of June 2021. In that range of time, we collected more than 9000 tweets from Indonesian users. Then, we also remove duplicate tweets in order to avoid overfitting. From those tweets, we can identify that its composition in Table IV.

TABLE IV. DATA LABEL FOR CLASSIFICATION

Label	Number
Positives	1767
Neutral	579
Negatives	1211
Unrelated	6303
Total Related Data	3557

After data gathering and labeling is finished, the next step is we throw away unrelated tweets. It is because most of the unrelated data that have been gathered mainly asked about how others' opinion about the policy. Since we wanted to build a model that express their opinion related into the topic rather than asking for others' opinion about the topic, we decided to dropped all of the unrelated data. Therefore, the total data that will be processed are 3557 data, consisting only Neutral, Negatives, and Positives labelled data.

B. Experiment Results and Evaluation

The features from the gathered data were extracted using TF-IDF for "tweet" and StandardScaler for "count". After that, we use cross validation with ShuffleSplit to divide the data into 5 parts that will be tested with each other. Then, it will be processed using Linear SVM, MLP with 10 nodes in each of 10 hidden layers, adam optimizer, and 100 iterations (MLP A), and MLP with 15 nodes in each of 10 layers, adam optimizer, and 100 iterations. After that, we evaluate the result by determine its scoring: accuracy, precision, recall, and F1 of the model. All of the metrics scoring use weighted value because there is data imbalance for each class. Then, we use the average score from each of the cross validation using ShufleSplit iterations. Some of the experiments or scenarios results can be seen in Table V and Table VI.

TABLE V. UNIGRAM FEATURE EXTRACTION RESULT

Feature(s)	Metrics	Class		
		SVM	MLP-A	MLP-B
Only Tweet	Accuracy	**68.5%**	**65.1%**	**63.3%**
	Precision	68.8%	64.6%	63.0%
	Recall	68.5%	65.1%	63.3%
	F1-Score	**67.0%**	**64.1%**	**63.1%**
Tweet + Count	Accuracy	**68.7%**	**65.0%**	**62.9%**
	Precision	69.2%	64.6%	63.0%
	Recall	68.7%	65.0%	62.9%
	F1-Score	**67.3%**	**64.6%**	**62.0%**
Tweet + Hours	Accuracy	**68.7%**	**63.4%**	**69.1%**
	Precision	69.2%	63.3%	68.6%
	Recall	68.7%	63.4%	69.1%
	F1-Score	**67.3%**	**64.1%**	**67.8%**
Tweet + Count + Hours	Accuracy	**68.7%**	**65.1%**	**68.7%**
	Precision	69.1%	64.6%	69.1%
	Recall	68.7%	65.1%	68.9%
	F1-Score	**67.3%**	**64.1%**	**67.8%**

TABLE VI. BIGRAM FEATURE EXTRACTION RESULT

Feature(s)	Metrics	Class		
		SVM	MLP-A	MLP-B
Only Tweet	Accuracy	**65.4 %**	**61.6%**	**51.4%**
	Precision	70.0%	61.8%	65.5%
	Recall	65.4%	61.6%	51.4%
	F1-Score	**62.3%**	**60.9%**	**51.1%**
Tweet + Count	Accuracy	**65.7%**	**62.9%**	**62.2%**
	Precision	70.3%	63.0%	61.7%
	Recall	65.7%	62.9%	62.2%
	F1-Score	**62.7%**	**62.0%**	**61.4%**
Tweet + Hours	Accuracy	**65.5%**	**63.4%**	**62.1%**
	Precision	70.3%	62.7%	61.4%
	Recall	65.5%	63.4%	62.1%
	F1-Score	**62.5%**	**62.4%**	**61.4%**
Tweet + Count + Hours	Accuracy	**65.7%**	**61.6%**	**62.2%**
	Precision	70.3%	61.8%	61.7%
	Recall	65.7%	61.6%	62.2%
	F1-Score	**62.6%**	**60.9%**	**61.4%**

To answer RQ1, our research use SVM and ANN with MLP that can be seen in Table V and Table VI. In unigram model (Table V), SVM with only tweet as the feature have 68.5% accuracy. After we add more features, its accuracy tends to slightly increased into 68.7%. Similar result also can be seen for MLP-B. Its accuracy slightly increased from 63.3% and varies into 62.9%, 69.1%, and 68.1%. In the other hand, MLP-A accuracy tends to slightly decreased. It went from 65.1% into 65.0%, 63.4%, and 65.1%.

In bigram model (Table VI), SVM with only tweet as the feature have 65.4% accuracy. After we add more features, its accuracy tends to slightly increased into 65.5% and 65.7%. The experiment is the same for MLP-B. Its accuracy increased from 51.4% into 62.2%, 62.1%, and 62.2%. Surprisingly, MLP-A accuracy also tends to increased slightly. It went from 61.6% into 62.9%, 63.4%, and 61.6%.

From the Table V and Table V, we can say that the models that used unigram generally produced better results than the models that used bigram in this study. This might be related to the sparsity of the data. A tweet in Twitter is limited to 140 characters, thus would also limit the words Twitter users can express. Because of that, more token can be created with unigram compared to bigram, which resulted in more distinct features within the data. With the addition of more features as shown in Table VI, the accuracy for bigram increased significantly up to around 11% compared to the model that only use tweet (from 51.4% to 62.2% using MLP B), even though the results are not the same with other algorithm. Therefore, using unigram as a text representation model for Twitter text processing is more suitable than other n-grams text representation, but we can add more features and use suitable Neural Network algorithm to overcome the problem using two or more text representation.

To answer RQ2, we tried to analyze text and other additional features to be added into tweet data. By adding additional features, we can produce slightly higher accuracy as shown in Table V and Table VI, where the compared to the model with no feature added. For this experiment, the best result is achieved by using unigram with tweet and count as the features of and MLP-B as its machine learning technique. Therefore, we conclude that any additional features proposed in this study can be useful to increase the metrics of the proposed models, even if the change is only minimal.

For the next part of this study, we will look at the distribution of the tweet that used in this study. The data is divided into Positives, Neutral, and Negatives. The result of this process can be seen in Table VII and Table VIII.

TABLE VII. TOP 3 TWEET FROM EVERY LABEL

TOP 3 TWEETS			
LABEL	NO	TWEETS	FREQUENCY
Positives	1	PRO. Harus emng harus bangt dilaksanakan offline spy proses transfer ilmu dari pengajar ke yang diajar lebih efektif. Sekolah online nge buat siswa kejar nilai bukan ilmu. Kuliah susah banget cok klo online. Apalagi jurusan yang bnyk praktikum hadeeeh..	84
	2	sezuzurnya mau kuliah offline supaya bisa secepatnya cabut dari rumah	52
	3	Dan di semester 4 ini aku udah mulai pasrah, kehilangan semangat kuliah, mulai capek online. Pengen kuliah	20

TABLE VIII. TOP 3 TWEET FROM EVERY LABEL(CONT.)

TOP 3 TWEETS			
LABEL	NO	TWEETS	FREQUENCY
Neutral	1	Jika pun maju dengan opsi sekolah tatap muka, maka selain 3M, bukalah semua pintu dan jendela setiap sekolah.	39
	2	Saya memandang, arahan Presiden Jokowi tentang sekolah tatap muka sebagai jalan tengah. Mulai dari pembatasan jam belajar…	18
	3	netral sih tbh, soalnya emang sekolah online tuh gak efektif banget mana guru semena-mena aja. tapi kalo offline…	12
Negatives	1	4 risiko Covid-19 pada anak saat sekolah tatap muka dimulai lagi: 1. Risiko sakit berat & kematian (terutama bila ada komplikasi…	120
	2	Dan positive rate hari ini di atas 13%. Lalu ada WNA dari India yg dateng, terus Wapres minta santri boleh mudik, tempat wisata…	10
	3	Mengakhiri thread ini, saya harap orang tua siswa sudah tahu dan paham mengenai risiko-risiko ini sebelum mengizinkan anak	7

From Table VII, we can see most common tweet from each label. All of them are retweets from one of the influencers in Indonesia. In Positives label, people urge the policy to be embraced because of several reasons. In the most popular tweet, it implies that face-to-face learning is important for learning effectivity. In the other hand, the second most popular tweet implies that student in university prefer to be far away from home. But in the third most popular tweet, it implies that the university student feels bored by the online learning program after more than a year.

In Table VIII Neutral label, people tend to find some alternatives that have been embraced by the Government. In the most popular tweet, they agree to take extra measure for Covid-19 prevention by open up the classroom doors and windows other than 3M measure (wash hand, use mask, distancing). The second most popular tweet also have similar note. They agree to the President policy for new face-to-face learning. In the third most popular tweet, they simply express their neutrality by giving positives and negatives side of face-to-face learning.

In Table VIII Negatives label, people tend to tweet about the risk of face-to-face learning amid the pandemic. In the most popular tweet, it was a retweet that stated about the risks of Covid-19 to children. In the second most popular tweet, it was stated the increasing of India-variant of Covid-19. In the third most popular tweet, it was the final thread, or the continuity of the most popular tweet.

V. CONCLUSION

People have different opinions and post it on social media. Many of them contain emotion and information in every post they made. With those emotion, we can have some insight about every topic discussed on social media. With the plan of face-to-face learning policy, many Indonesians have different opinion about the topic. By using machine learning approach, we can determine where they stand, positive, negative, or neutral about the topic. From this study we can conclude that adding features can affects the result of the experiment. It slightly increases the accuracy, precision, recall and F1-Score compared to only using tweets as the processed data, even if it was only slight improvement.

This study also found that there are many opinions in form of tweets about the face-to-face learning system that have been proposed by The Government. By dividing these tweets into 3 major group, we can see what the users mostly talk about. From Positives label, we know that the people of Indonesia thinks that face-to-face learning is important for knowledge transfer for student. Other than that, the students simply feel bored by study at home.

In the other hand (Negatives label), there are concern about the safety of the student. They are afraid for their children to be exposed to Covid-19. Between these two major groups (Neutral labels), they expect some solution from the Government to overcome the problem of this topic. Therefore, the Government needs to find the best alternative for education program in Indonesia.

ACKNOWLEDGEMENT

Special thanks to Dr. Indra Budi (Universitas Indonesia) as our supervisor who have been teaching and discussing with us about sentiment analysis. We would also like to thank Ibid Athoillah, Intan Sari Putri, Ferry Wahyu Irzadiawan and Dewa Made Oka Purnama Atmaja that have helped us to labeling the data. Without their help, this study would take a longer time to finish.

REFERENCES

[1] World Health Organization, "Novel Coronavirus (2019-nCoV) SITUATION REPORT - 2," 2020.

[2] Bloomberg.com, "Indonesia Bans Annual Holiday Exodus to Combat Virus Spread," 2020. https://www.bloomberg.com/news/articles/2020-04-21/indonesia-bans-mass-travel-ahead-of-eid-festival-to-combat-virus (accessed Jun. 21, 2021).

[3] Kementerian Penddidikan dan Kebudayaan, "Panduan: Pembelajaran Jarak Jauh," *kemdikbud.go.id*, 2020. https://bersamahadapikorona.kemdikbud.go.id/panduan-pembelajaran-jarak-jauh/ (accessed Mar. 15, 2021).

[4] H. Yulia, "Online Learning to Prevent the Spread of Pandemic Corona Virus in Indonesia," *ETERNAL (English Teach. Journal)*, vol. 11, no. 1, pp. 48–56, 2020, doi: 10.26877/eternal.v11i1.6068.

[5] Rasmitadila *et al.*, "The perceptions of primary school teachers of online learning during the covid-19 pandemic period: A case study in Indonesia," *J. Ethn. Cult. Stud.*, vol. 7, no. 2, pp. 90–109, 2020, doi: 10.29333/ejecs/388.

[6] The Jakarta Post, "Majority of schools have reopened limited classes, says survey," 2021. https://www.thejakartapost.com/adv/2021/06/21/majority-of-schools-have-reopened-limited-classes-says-survey.html (accessed Jun. 21, 2021).

[7] The Jakarta Post, "School reopening plan in doubt amid COVID-19 surge," 2021. https://www.thejakartapost.com/paper/2021/06/22/school-reopening-plan-in-doubt-amid-covid-19-surge.html (accessed Jun. 21, 2021).

[8] The Jakarta Post, "Education Can't Wait," 2021. https://www.thejakartapost.com/academia/2021/05/02/education-cant-wait.html (accessed Jun. 21, 2021).

[9] K. Z. Aung and N. N. Myo, "Sentiment analysis of students' comment using lexicon based approach," *Proc. - 16th IEEE/ACIS Int. Conf. Comput. Inf. Sci. ICIS 2017*, pp. 149–154, 2017, doi: 10.1109/ICIS.2017.7959985.

[10] E. Cambria, S. Poria, A. Gelbukh, and M. Thelwall, "Sentiment Analysis Is a Big Suitcase," *IEEE Intell. Syst.*, vol. 32, no. 6, pp. 74–80, 2017, doi: 10.1109/MIS.2017.4531228.

[11] V. Kecman, "Support Vector Machines – An Introduction 1 Basics of Learning from Data," *StudFuzz*, vol. 177, pp. 1–47, 2005, [Online]. Available: http://mplbci.ekb.eg/MuseProxyID=1104/MuseSessionID=08102g99u/MuseProtocol=https/MuseHost=link.springer.com/MusePath/content/pdf/10.1007%2F10984697_1.pdf.

[12] R. Moraes, J. F. Valiati, and W. P. Gavião Neto, "Document-level sentiment classification: An empirical comparison between SVM and ANN," *Expert Syst. Appl.*, vol. 40, no. 2, pp. 621–633, 2013, doi: 10.1016/j.eswa.2012.07.059.

[13] J. A. Anderson, "Book Reviews An Introduction to Neural Networks The Engine of Reason , the Seat of the Soul : A Philosophical Journey into the Brain," *Neuroscience*, pp. 383–385, 1995.

[14] L. Rajabion, K. Wakil, A. Badfar, S. Mojtabavi Naeini, and B. Zareie, "A new model for assessing the impact of ICT and digital knowledge on students' thoughts and beliefs," *J. Eng. Des. Technol.*, vol. 17, no. 5, pp. 943–959, 2019, doi: 10.1108/JEDT-01-2019-0008.

[15] T. Hastie, R. Tibshirani, and J. H. Friedman, *The elements of statistical learning: data mining, inference, and prediction*, vol. 2. 2009.

[16] D. Mowery, C. Bryan, and M. Conway, "Feature studies to inform the classification of depressive symptoms from Twitter data for population health," *arXiv*, pp. 0–4, 2017.

[17] E. A. Sukma, A. N. Hidayanto, A. I. Pandesenda, A. N. Yahya, P. Widharto, and U. Rahardja, "Sentiment Analysis of the New Indonesian Government Policy (Omnibus Law) on Social Media Twitter," *Proc. - 2nd Int. Conf. Informatics, Multimedia, Cyber, Inf. Syst. ICIMCIS 2020*, pp. 153–158, 2020, doi: 10.1109/ICIMCIS51567.2020.9354287.

[18] Z. Kamisli Ozturk, Z. I. Erzurum Cicek, and Z. Ergul, "Sentiment analysis: An application to Anadolu University," *Acta Phys. Pol. A*, vol. 132, no. 3, pp. 753–755, 2017, doi: 10.12693/APhysPolA.132.753.

[19] S. Shahnawaz and P. Astya, "Sentiment analysis: Approaches and open issues," *Proceeding - IEEE Int. Conf. Comput. Commun. Autom. ICCCA 2017*, vol. 2017-Janua, pp. 154–158, 2017, doi: 10.1109/CCAA.2017.8229791.

[20] Y. Arakawa, K. Yasumoto, K. Pattamasiriwat, and T. Mizumoto, "Improving recognition accuracy for activities of daily living by adding time and area related features," *2017 10th Int. Conf. Mob. Comput. Ubiquitous Network, ICMU 2017*, vol. 2018-Janua, pp. 1–6, 2018, doi: 10.23919/ICMU.2017.8330104.