

PAPER • OPEN ACCESS

Online learning sentiment analysis during the covid-19 Indonesia pandemic using twitter data

To cite this article: Syafrida Hafni Sahir *et al* 2021 *IOP Conf. Ser.: Mater. Sci. Eng.* **1156** 012011

View the [article online](#) for updates and enhancements.

You may also like

- [Online Learning in Faculty of Mathematics and Natural Sciences, State University of Medan: Lecturers' Preferences for Online Learning Media During the Covid-19 Pandemic](#)
Herbert Sipahutar, Fauziyah Harahap, Mukhtar et al.
- [Problems of students in following the online learning process in the covid-19 pandemic](#)
A Akmam, R Hidayat, F Mufit et al.
- [Online learning mediated by social teaching platforms: an experience from a flipped undergraduate physics course in renewable energies](#)
Matteo Tuveri, Alessia Zurru, Daniela Fadda et al.



ECS The Electrochemical Society
Advancing solid state & electrochemical science & technology

247th ECS Meeting
Montréal, Canada
May 18-22, 2025
Palais des Congrès de Montréal

Showcase your science!

ECS UNITED

Abstract submission deadline extended: December 20

Online learning sentiment analysis during the covid-19 Indonesia pandemic using twitter data

Syafrida Hafni Sahir¹, Raden Sri Ayu Ramadhana², Muhammad Fauzi Romadhon Marpaung², Shabrina Rasyid Munthe² and Ronal Watrianthos^{2*}

¹Department of Management, Universitas Medan Area, Indonesia

²Department of Informatics, Universitas Al Washliyah Labuhanbatu, Indonesia

* ronal.watrianthos@gmail.com

Abstract. In the last week of September 2020, Covid-19 in Indonesia has infected more than 252,000 people. The spread of the virus through physical contact forces all countries to use social distance and physical distance to reduce interactions. On the educational side, the Covid-19 pandemic has had an impact on the prevention of face-to-face teaching and learning activities. All educational institutions must stop the learning process and replace it with online learning. Nadiem Makarim, Minister of Education and Culture of the Republic of Indonesia, then affirmed this online learning by issuing a policy on online learning in case of an emergency spread of the Covid-19 virus. However, this online learning has provoked debate in the community over the preparation of the technology. The goal of this study was to analyze public opinion for online learning during the Covid-19 pandemic in October 2020. The research carried out document-based text mining and feelings on Twitter data analyzed using the Naïve Bayes method. Analysis in that period found 25 percent positive sentiment, 74 percent negative sentiment, and 1 percent neutral sentiment. Several tweets showed that the words 'stress' and 'covid' were the most commonly spoken in the month.

1. Introduction

The World Health Organization (WHO) declared the COVID-19 virus a global pandemic on March 11, 2020 [1]. WHO reports 35 million people infected with Covid-19 and 1 million more people died in early October 2020. Meanwhile, Indonesia recorded 307 thousand people confirmed positive with the death toll has reached 11,250 people [2]. The rapid spread of the virus through physical contact forces all countries to apply social distancing and physical distancing to reduce interactions. The main factor for transmission through droplets released during talking, coughing, or sneezing makes social distancing the most widely adopted strategy when there is no vaccine [3].

Indonesia's President Jokowi has issued a Large-Scale Social Restriction (PSBB) policy [4]. Social restrictions affect public spaces, including offices, schools, or campuses. In education, this strategy prohibits face-to-face learning and forces all educational institutions to stop the learning process and replace it with online learning [5]. In education, difficulties emerge, such as the unpreparedness of technologies, the media used, or the psychological aspects of students. In the meantime, the future will be unequal education between populations and regions in Indonesia in the long term [6].

The plan to tackle the pandemic of social constraints pushes all educational institutions to embrace online learning [5]. The Ministry of Education and Culture of the Republic of Indonesia Nadiem Makarim underscored the policy of online learning during the pandemic by issuing the Learning From



Home (BDR) policy [7]. This policy requires the use of the Internet network through the intermediary of smartphones, gadgets, computers, and applications as a substitute for face-to-face use.

Online learning, which is called a strategy, is then problematic because it has to be adapted. Infrastructure gaps, the efficiency of connections, devices used, and internet prices are the biggest hurdles. The abrupt transition from face-to-face learning to large-scale online learning has given rise to a diverse variety of responses in public. This study aims to analyze online education public sentiment during the October 2020 pandemic of Indonesia COVID-19. This research uses the keywords 'online learning, lecture, learning,' and the hashtag '#BelajarDariRumah' which are filtered with the keyword 'online' and 'home' on Indonesian tweets which are scrambled from Twitter and processed using the Naïve Bayes algorithm to produce sentiment analysis on online learning acceptance.

2. Methodology

2.1 Sentiment Analysis

Analysis of sentiment is a method to recognize opinion and classify text in a document/phrase to identify groups as positive, negative, or neutral [8]. Today, academics research is commonly using emotion analysis as a branch of computer science research. Social networks, such as Twitter, are widely used in sentiment analysis to classify public views [9].

It is also possible to equate sentiment analysis with opinion mining since it relies on viewpoints that are favorable or negative [10]. Data mining is applied to analyze, process, and retrieve textual data within an object in sentiment analysis, such as utilities, goods, persons, phenomena, or other subjects. The research process can involve text reviews, forums, tweets, or blogs, with pre-processing data covering tokenization, stopword, deletion, stemming, identification of emotions, and classification of sentiments [11].

2.2 Text Mining

Text mining is known to produce information from a set of documents. Text Mining can generate information by processing, grouping, and analyzing large volumes of unstructured data [12]. Data mining is to extract valuable information from a variety of documents with unstructured text data sources [13]. The process of collecting data in text mining can produce a sensing analysis that identifies an emotionally positive or negative statement [14].

2.3 Naïve Bayes Method

The Naïve Bayes method is a text mining classification method used for the study of emotion. In terms of data consistency and measurement classification, this method is technically reasonable [15]. In the Twitter classification method, Naïve Bayes uses a variety of approaches, including Unigram Naïve Bayes, Multinomial Naïve Bayes, and Maximum Entropy Classification [16]. The main feature of the Naïve Bayes classification is the firm hypothesis of each situation or event [17]. Calculation of the probability groups of Naïve Bayes uses the Bayes algorithm method using the following equation:

$$P(c|z) = \frac{p(z|c) \times p(c)}{p(z)} \quad (1)$$

Equation (1) shows that Y is a specific class, X is data in the unknown classed, while P (Y | X) is the probability of the hypothesis based on conditions, while P (Y) and P (X | Y) is the previous probabilities of the class based on hypothetical, while P (X) is the probability Y [18]. In the Naïve Bayes classification, equation 1 is further developed into equation (2) as follows:

$$P(Y|X_1, X_2, \dots, X_n) = \frac{P(X_1, X_2, \dots, X_n|Y)P(Y)}{P(X_1, X_2, \dots, X_n)} = \frac{P(X_1|Y)P(X_2|Y) \dots P(X_n|Y)P(Y)}{P(X_1, X_2, \dots, X_n)} \quad (2)$$

Where $P(Y|X_1, X_2, \dots, X_n)$ is the result of the calculation of all posterior probabilities at the value of X for all Variable y, the Naïve Bayes classification will make predictions dependent on the maximal likelihood of the posterior Laplace probability shown in equation (3) below, where c is the number of Y values [21]:

$$P(X_i|Y) = \frac{N_{ic}+1}{N_c+c} \quad (3)$$

Figure 1 shows an opinion mining method using the Naïve Bayes process on Twitter. The process of crawling twitter data in this analysis is to provide keywords at a certain period. Labeling is applied as data is collected to determine emotions. The next step is the pre-processing and systematic transformation of data collection. At this point, a cleaning process is performed out to minimize noise and delete stopwords to eliminate useless words such as 'I,' 'and' or ' ' [5]. The tokenization method is used to define words and split sentences into space-based phrases and punctuation marks. The last step of pre-processing is to convert the affixes to simple phrases. The third stage of opinion mining is a feature extract intended to ease the classification of Naïve Bayes. This stage produces a model which is used to show the consistency of the effects of the classification [19].

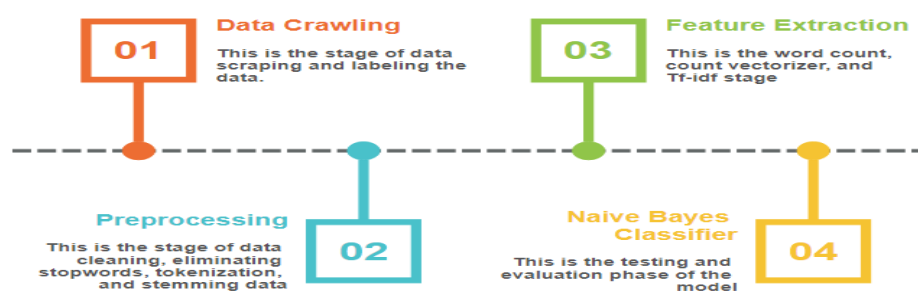


Figure 1. The Naïve Bayes Process for Classifying

The results of the crawling tweets are stored in a CSV file and split into two datasets, namely training data and test data. Positive and negative tweets are categorized on a label list. The Naïve Bayes method will be used at the stages of sentiment classification and interpretation of the result of the sentiment analysis.

3. Result and Discussion

We use the terms of online studying, lecture, research, and hashtag # BelajarDariRumah as keywords for Indonesian tweets. The data crawling method using the public Twitter stream of Drone Emprit Academic was able to capture 159,045 tweets linked to the chosen keywords shown in Table 1 [20].

Table 1. Crawling Data

@aliefriendshew	Sekolah online : tambah bego Sekolah offline : meresahkan Yaallah capek bgt akutu
@nesshlyn	Buat yg kelas 9 sama 12 ngerasa ngga? Udah nyaman sekolah online, kalau offline pun udah males duluan. Tapi kepikiran pelajaran sama sekali ngga ngerti takut buat kedepanya plonga-plongo
@xxzaww	saya lebih menyukai sekolah offline dari pada online,terimakasih
@schfess	Sch! Semangat buat kamu yg tugasnya masih numpuk walopun libur!! kita semua disini sama-sama berjuang sekolah online jadi jgn ngerasa sendirian/putus asa ya!! https://t.co/IBf2170OFM
@dulce_maria_lrs	Gue pribadi cukup miris dengan guru yang kasih tugas berlebihan untuk murid, dengan alasan sekolah daring. Bayangin aja,sekolah online by zoom,gru jelasin, dan dikasih tugas setumpuk lagi, itu baru 1 mapel, sedang sehari 3-4 mapel, lu bayangin itu otak kerjanya kayak apa+

The process of tokenization and data conversion occurs during preprocessing to remove noise and clarify features. The data cleansing stage removes unnecessary attributes such as URL, mention, username, RT, hashtag, or punctuation and class assignments. The preprocessing results were grouped

to get the most word frequency and visualized into positive and negative word cloud shown in Figure 2.



Figure 2. Wordclouds Positive and Negative

The sentiment analysis model is built based on training data that has been prepared and tested to determine sentiment towards online learning. The text analyzed in this study is the text in Indonesian tweets and uses Indonesian vocabulary [21]. The results of the sentiment test data analysis on the training data could be in Table 2 for five sample data.

Table 2. Test Data Analytical Results

Sentiment	Confident (Positive)	Confident (Neutral)	Confident (Negative)	Text (in Bahasa)
Positive	0.9856	0.0071	0.0073	Semangat buat yang lagi sekolah online kuliah online, ngerjain tugas, magang, kerja dari rumah maupun harus keluar, dan kegiatan apapun. Tetap hati2 dan jaga kesehatan ya!
Negative	0.0072	0.0070	0.9860	kuliah online bikin weekend jadi ga punya harga diri lagi like, sekarang mau hari kerja atau akhir pekan ga ada bedanya, full meeting, full tugas, full deadline semua stay hydrated all
Negative	0.0076	0.0074	0.9877	Kendala kuliah online gue tiap bimbingan gue paling sebel kalau pas lagi asistensi terus tiba2 emak gue buka pintu triak nyuruh cuci piring pdhl lagi vidio call sama dospem
Negative	0.0071	0.0073	0.9871	lebih ke capek, rutinitas kuliah online kerasa kalau gitu gitu aja soalnya. kuliah-nugas-kuliah-nugas dan itu harus ngadep laptop terus. lama lama jenuh terus capek aja rasanya. capek mata sama capek badan juga
Positive	0.0989	0.0070	0.0075	jaman sekarang kalo mau belajar mah udah enak ya, bisa belajar online diajarin tutor yang ramah nan ceria, jaman w dulu belajar di rumah diajarin emak, digembleng ampe nangis2

Just 25 percent had a positive sentiment based on the findings of the study tested to get a sense of online learning on Twitter in the October period. At 75 percent, negative opinion dominates, and the remaining 1 percent is neutral.

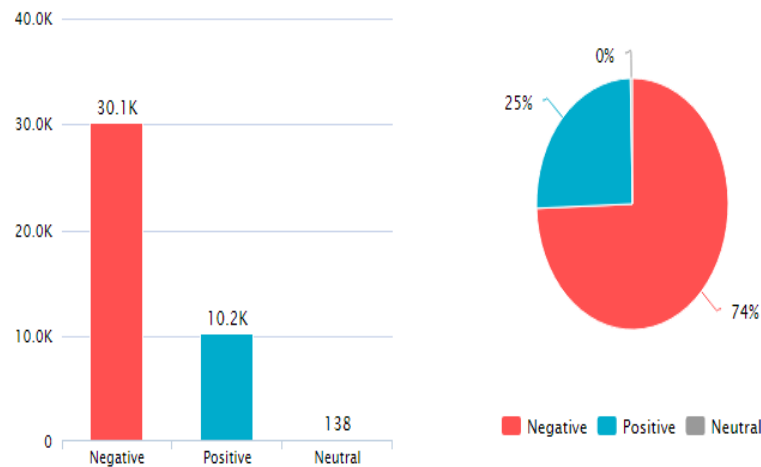


Figure 3. Measurement of sentiment with Naïve Bayes.

Figure 3 shows the sentiment measurement results with more than 30 thousand tweets being negative, more than 10 thousand tweets giving positive results, and the remaining 138 tweets were considered neutral during the month. Based on these results, online learning has not been maximally applied in Indonesia during the pandemic, as seen from the high public disappointment.

4. Conclusion

The Naïve Bayes algorithm for sentiment classification applies to this analysis. Sentiment analysis in this research uses Twitter data with the keywords 'online learning' 'school' 'studying' and the hashtag #BelajarDariRumah filtered with the keywords 'online' and 'home' on Indonesian tweets in October 2020. Analysis in that period found 25 percent positive sentiment, 74 percent negative sentiment, and 1 percent neutral sentiment. People's frustration with online learning produced negative perceptions. Several tweets revealed discontent with the words 'stress' and 'covid' having the highest communication level throughout the month. In future experiments, the use of various algorithms is likely to provide more reliable performance.

References

- [1] D. Cucinotta and M. Vanelli, "WHO declares COVID-19 a pandemic," *Acta Biomed.*, vol. 91, no. 1, pp. 157–160, 2020, doi: 10.23750/abm.v91i1.9397.
- [2] WHO, "WHO Coronavirus Disease (COVID-19) Dashboard," *WHO Health Emergency Dashboard*, 2020. <https://covid19.who.int/> (accessed Oct. 07, 2020).
- [3] D. M. Dave, A. I. Friedson, K. Matsuzawa, J. J. Sabia, and S. Safford, "Black lives matter protests, social distancing, and COVID-19," *NBER Work. Pap. Ser.*, 2020.
- [4] A. Amindoni, "Virus corona: Presiden Jokowi pilih 'pembatasan sosial dalam skala besar', warga mulai sortir pendatang," *BBC News Indonesia*, 2020. <https://www.bbc.com/indonesia/indonesia-52059236> (accessed Oct. 07, 2020).
- [5] R. Watrianthos, "Analisis Pembelajaran Daring di Era Pandemi Covid-19," in *Merdeka Kreatif di Era Pandemi Covid-19: Suatu Pengantar*, Medan: Green Press, 2020, p. 55.
- [6] R. H. Syah, "Dampak Covid-19 pada Pendidikan di Indonesia: Sekolah, Keterampilan, dan Proses Pembelajaran," *SALAM J. Sos. dan Budaya Syar-i*, vol. 7, no. 5, Apr. 2020, doi: 10.15408/sjsbs.v7i5.15314.
- [7] Kumparan, "4 Kebijakan Nadiem Makarim soal Proses Belajar dari Rumah Selama Pandemi Corona," *kumparan.com*, 2020. <https://kumparan.com/kumparanmom/4-kebijakan-nadiem-makarim-soal-proses-belajar-dari-rumah-selama-pandemi-corona-1t5naOVW9MB/full> (accessed Oct. 07, 2020).
- [8] D. A. Ramadhan and E. B. Setiawan, "Analisis Sentimen Program Acara Di Sctv Pada Twitter

- Menggunakan Metode Naive Bayes Dan Support Vector Machine,” in *e-Proceeding of Engineering*, 2019, pp. 9376–9743.
- [9] R. Ferryawan, Kusrini, and F. W. Wibowo, “Analisis Sentimen Wisata Jawa Tengah Menggunakan Naïve Bayes,” *J. Inf. Politek. Indonusa Surakarta*, vol. 5, no. 3, pp. 55–60, 2019.
 - [10] C. B. Saputra, A. Muzakir, and D. Udariansyah, “Analisis Sentimen Masyarakat Terhadap #2019gantipresiden Berdasarkan Opini Dari Twitter Menggunakan Metode Naive Bayes Classifier,” in *Bina Darma Conference on Computer Science*, 2019, pp. 403–413.
 - [11] S. B. Bhonde and J.R. Prasad, “Sentiment Analysis-Methods, Applications and Challenges,” *International J. Electron. Commun. Comput. Eng.*, vol. 6, no. 6, pp. 634–640, 2015.
 - [12] A. Sholihin, Haviluddin, N. Puspitasari, M. Wati, and Islamiyah, “Analisis Penyakit Difteri Berbasis Twitter Menggunakan Algoritma Naïve Bayes,” *SAKTI – Sains, Apl. Komputasi dan Teknol. Inf.*, vol. 1, no. 1, pp. 7–15, 2019.
 - [13] G. E. I. Kambey, Rizal Sengkey, and Agustinus Jacobus, “Penerapan Clustering pada Aplikasi Pendeteksi Kemiripan Dokumen Teks Bahasa Indonesia,” *J. Tek. Inform.*, vol. 15, no. 2, pp. 75–82, 2020.
 - [14] X. Zhou, X. Tao, and Z. Yang, “Sentiment Analysis on Tweets for Social Events,” in *IEEE 17th International Conference on Computer Supported Cooperative Work in Design*, 2013, pp. 557–562.
 - [15] J. M and V. H, “Opinion Mining For Sentiment Data Classification,” *Int. J. Res. Inf. Technol.*, vol. 3, no. 1, pp. 1–13, 2014.
 - [16] R. P and M. M, “Sentiment Analysis of User Generated Twitter Updates using Various Classification,” 2009.
 - [17] N. Rochmawati and S. C. Wibawa, “Opinion Analysis on Rohingya using Twitter Data,” *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 336, no. 1, 2018, doi: 10.1088/1757-899X/336/1/012013.
 - [18] Syahriani, A. A. Yana, and T. Santoso, “Sentiment analysis of facebook comments on indonesian presidential candidates using the naïve bayes method,” in *Journal of Physics: Conference Series*, Nov. 2020, vol. 1641, p. 012012, doi: 10.1088/1742-6596/1641/1/012012.
 - [19] R. Watrianthos, S. Suryadi, D. Irmayani, M. Nasution, and E. F. S. Simanjorang, “Sentiment Analysis Of Traveloka App Using Naïve Bayes Classifier Method,” *Int. J. Sci. Technol. Res.*, vol. 8, no. 07, pp. 786–788, 2019.
 - [20] I. Fahmi, “Drone Emprit Academic: Software for social media monitoring and analytics,” *dea*, 2020. academic.droneemprit.id (accessed Oct. 31, 2020).
 - [21] D. R. Lazuardi, T. A. Munandar, H. Harsiti, Z. Mutaqin, and R. N. Hays, “Sentiment analysis of public opinions on the welfare of honorary educators using Naive Bayes,” *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 830, p. 032018, May 2020, doi: 10.1088/1757-899X/830/3/032018.