RESEARCH ARTICLE | MARCH 26 2024

Sentiment analysis regarding the impact of Covid-19 on education in Indonesia with the naïve bayes classifier FREE

Christian Sri Kusuma Aditya

AIP Conf. Proc. 2927, 060022 (2024)
https://doi.org/10.1063/5.0194868

Sentiment Analysis Regarding the Impact of Covid-19 on Education in Indonesia with The Naïve Bayes Classifier

Christian Sri Kusuma Aditya[a)]

Department of Informatics, Universitas Muhammadiyah Malang, Malang, Indonesia

[a)]Corresponding author: christianskaditya@umm.ac.id

Abstract. In December 2019, the new coronavirus SARS-CoV-2 unexpectedly produced the COVID-19 pandemic in China. The World Health Organization reports millions of verified cases and more than a hundred thousand confirmed deaths globally. As a result, during various outbreak-related occurrences, these social media platforms are exposed to and display a variety of perspectives, ideas, and feelings. Lately, there have been many tweets from the public regarding the dire conditions that have occurred in Indonesia since the COVID-19 outbreak, especially related to its impact on the education sector in Indonesia, which until now still often causes pros and cons in most surrounding communities. One of the effective classification methods for sentiment analysis methods is Nave Bayes. Naïve Bayes is applied by determining the appearance of sentiment contained in tweets. Before applying the Naïve Bayes method, data preprocessing and application of the TF-IDF method were carried out. From the research that has been done, the accuracy rate of the Naive Bayes Classifier algorithm is 0.696 with a procedure without the stemming stage and 0.705 with the stemming stage present. The presence or absence of the stemming process significantly impacts the classification's outcome; in this study, the classification's final value was slightly higher when the stemming step was used than when it was not.

Keywords: Sentiment Analysis, Covid-19, Education, Naive Bayes

INTRODUCTION

Talking about freedom in writing or expressing something on social media, there are not a few tweets or news that are being hotly discussed on Twitter social media and always quickly become trending. One of them is about the pandemic or outbreak that is being felt by the world, especially in Indonesia, and how the Indonesian people respond to the impact of COVID-19 on education in Indonesia.

Lately, there have been many tweets from the public regarding the dire conditions that have occurred in Indonesia since the COVID-19 outbreak, especially related to its impact on the education sector in Indonesia, which until now still often causes pros and cons in most surrounding communities.

As written in the article entitled "The Impact of Covid-19 on Education in Indonesia: Schools, Skills, and the Learning Process" [1] said that there are many variants of problems that hinder the effectiveness of learning during the COVID-19 pandemic, namely the method online include limited mastery of information technology by teachers and students, inadequate facilities and infrastructure, limited internet access, and lack of preparation in providing budget.

In contrast, [2], in the article entitled "Positive Impact of the Covid-19 Pandemic on the World of Education," said that the current learning method is a positive application where this method can motivate through difficult times to continue to achieve Indonesia's educational goals. More advanced ones such as triggering the acceleration of educational transformation, the emergence of online learning applications, the number of free online courses, the emergence of unlimited creativity, a collaboration between teachers and parents, the application of knowledge in the family, making teachers more "literate" and familiar with technology, the internet is a positive source of information, and parents can directly supervise students.

2nd International Conference on Technology, Informatics, and Engineering
AIP Conf. Proc. 2927, 060022-1–060022-7; https://doi.org/10.1063/5.0194868
Published under an exclusive license by AIP Publishing. 978-0-7354-4901-5/$30.00

This research was conducted using the Naïve Bayes Classifier algorithm. Based on several previous studies, the Naïve Bayes method in sentiment analysis has a reasonably high and effective result value even for extensive data in terms of accuracy, precision, or recall compared to several other methods.

In one of the studies conducted [3], it is known that the Naïve Bayes method has a higher value than KNN, one of which is an accuracy rate of 63.21 %, while the KNN method is 58.10%, and it is also found that the tendency of public opinion on Twitter tends to be positive, and this can be seen from the number of positive opinions of 610, negative of 488, which is supported by the results of precision testing in the Naïve Bayes method with more positive values higher than negative that is 66.40%: 58.94%

Research in [4] shows that the Naïve Bayes method gets the highest results, namely 86.53% on testing data and 94.08% on data training. Meanwhile, Lexicon Based got a score of 64.49% on testing data and 94.2% on training data

This research was conducted to provide an overview of the community or even the Indonesian government regarding how the response of the Indonesian people to government policies in dealing with this disaster, especially to the ongoing education in Indonesia, so that the resulting data can later be used as a source of information for the community and the Indonesian government to be better and wiser in running the education system in Indonesia during the COVID-19 pandemic.

METHODS

In this section, the methods used in this research are explained. **FIGURE 1** shows the stages of the proposed method.

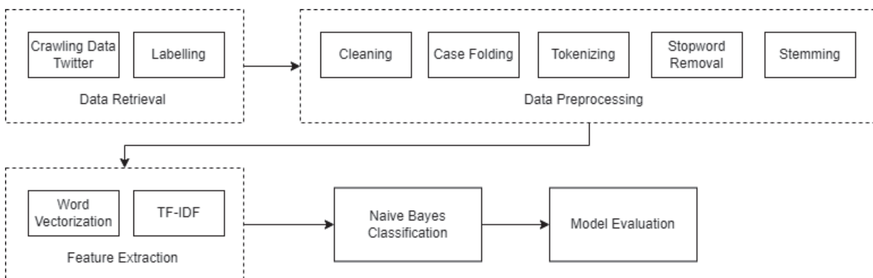

FIGURE 1. Step by step of the method used

The data used is a collection of tweets about the impact of COVID-19 on education in Indonesia, with several keywords such as online learning and Indonesian education. The data is obtained from Twitter social media with the help of services provided by Twitter for developers, namely the Twitter API (Application Programming Interface).

The data was obtained using a crawling technique with the Python programming language and 1000 tweets. **FIGURE 2** is an example of data resulting from the crawling process:

After data crawling is done, the next stage is preprocessing, which is the stage to clean the data so that the results obtained are quality information or documents that can facilitate the sentiment classification process. Several stages are carried out during data preprocessing. The first is cleaning, which aims to clean or remove punctuation marks, numbers, symbols, excess spaces, web links, and user names. The second stage is case folding, changing the capital letters in the sentence into lowercase letters. The third stage is tokenizing, the process of cutting words in a document into several word units or single words. The fourth stage is stopword removal, carried out to remove or delete words that have no meaning so that they are not needed in classification. The fifth stage is stemming, which removes affixes to get basic words or common words by removing prefixes, suffixes, insertions, and confixes (combinations of prefixes and suffixes). In this study, the stemming algorithm used is the Nazief & Adriani algorithm [5-7].

After the data preprocessing, the next step is labeling to determine opinions or views on the tweets. Labeling is divided into two classes, namely negative (0) and positive (1). **TABLE 1** is an example of data that has been labeled.

In this topic, the negative class (0) states that the tweets are words that show dissatisfaction or pessimism with the situation that is happening to the sustainability of the world of education when the Covid-19 pandemic is happening. In contrast, the positive class (1) shows an attitude that can be said to be optimistic or even neutral towards the sustainability of the world of education in Indonesia during the current pandemic.

Text
ðŸ'›Fidel`mulfandiðŸ"Œ, @kwonfiger Iya sama belajar daring bikin otak aku mau pecah rasanyaðŸ˜-
cek @schfess.id, Mau nanya dong buat kalian yg masih kelas 10-11 kalo mau PAS (penilaian akhir semester) kalo belajar suka dari mana?â€¦ https://t.co/Kg7FRrQdb3
.,@ismailfahmi ini stress karena belajar daring atau karena puasa ? ðŸ˜«
Akademi Farmasi Surabaya, "Hai Teman Farmasi! Belajar itu bisa kapan saja, dimana saja, dan lewat media apa saja. Sudah 1 Tahun lebih kita belajar daring.â€¦ https://t.co/74wQigP6Yo"
á´ÉªÉ´ ðŸ§šâ¤ï¸ðŸ¦‹, @flojnkyu Semenjak daring aku semangat belajar bahasa Jepang ðŸ˜ŒðŸ'
"@meyyuungi @iim_adelard mending belajar dulu, tugas daringmu masih numpuk noh, jangan urusin utang negara yang..â€¦ https://t.co/6rASejRIqW"
Mira Novita, "Meskipun bulan ramadhan,, saya tetap ngajar anak sekolah dari tingkat TK, SD, SMP, SMA Belajar daring maupun tatap muka untuk wilayah zona hijau."
kumparan, Asupan nutrisi ini penting agar anak-anak dapat menjalankan ibadah puasa serta aktivitas belajar daring dengan lebih..â€¦ https://t.co/W5OfuljLVu
ðŸ"¥ É†â‚ªÉ†â‚-É†â‚©É†â±¤ ðŸ"¥, "@NovalAssegaf @Aditya11278384 Masih aja belajar, lagi daring om ðŸ¤£ðŸ¤£"
@Hihihi_09, "Sebenernya belajar secara daring gak efektif karena apa?
Belajar tatap muka aku cuma seminggu. Minggu depan belajar daring.
Bulan Dua, Sesuatu yg sifat daring kayak kerjaan atau belajar online bukan gue banget. Maybe satu sisi efisien dapat ilmunya tapi emosional...â€¦ https://t.co/SyICyImtmm
"@ipb_menfess tidak, aKu ingin daring, hemat dompet, efford waktu, dll. + ada recordingnya (bisa belajar ulang)"
CEK PINNED DEH €"@SBMPTNFESS âœ¨, "selama daring, belajar pake hp terus ngerasa minus mata nambah gak sih? ini gimana caranya biar minusnya turun? sâ€¦ https://t.co/6NGyJr6zKY"
@Nakalkali69,@thewindNcloud Alangkah baiknya hpnya disumbangkan untuk adek adek yang sekolah daring...mereka kesulitan untuk..â€¦ https://t.co/ir3BjUayKj
@Urii,Diliburkan (belajar di rumah via daring)
BATAM TODAY, "Kasus Covid-19 Meningkat, Proses Belajar Mengajar di Anambas Dilakukan Secara Daring https://t.co/qsaSIaGS2b"

FIGURE 2. Data Crawl Results

The following process is the creation of features that are carried out to facilitate data classification, namely by performing feature extraction, and is done using TF-IDF (Term Frequency - Inverse Document Frequency) illustrated in **TABLE 2** [8, 9].

TABLE 1. Data Labeling

Tweet	Clean Text	Label
ðŸ"›Fidel`mulfandiðŸ"Œ, @kwonfiger Iya sama belajar daring bikin otak aku mau pecah rasanyaðŸ˜	*Fidel mulfandi Iya sama belajar daring bikin otak aku mau pecah rasanya*	0
cek @schfess.id, Mau nanya dong buat kalian yg masih kelas 10-11 kalo mau PAS (penilaian akhir semester) kalo belajar suka dari mana?â€¦https://t.co/Kg7FRrQdb3	*cek id Mau nanya dong buat kalian yg masih kelas kalo mau PAS penilaian akhir semester kalo belajar suka dari mana*	1
.,@ismailfahmi ini stress karena belajar daring atau karena puasa ? ðŸ˜«	*ini stress karena belajar daring atau karena puasa*	0
á´ ÉªÉ´ ðŸ§šâ□¤ï¸□ðŸ¦‹, @flojnkyu Semenjak daring aku semangat belajar bahasa Jepang ðŸ˜ŒðŸ"•	*Semenjak daring aku semangat belajar bahasa Jepang*	1

Several formulas can be used, including binary TF, which only pays attention to whether a term exists or not in the document. If it exists, it is assigned a value of one (1). Otherwise, it is assigned a value of zero (0). TF raw, where the weighting is given based on the number of term occurrences in the document. For example, if it occurs five (5) times, then the word will be worth five (5) [9].

Inverse Document Frequency (IDF) is a calculation where the term is widely distributed in the collection of documents concerned. The IDF shows the relationship between the availability of a term in all documents. The fewer the number of documents containing the term in question, the greater the IDF value. IDF is calculated using the formula as in **equation 1** [10].

$$IDF_j = \log\left(\frac{D}{df_j}\right) \tag{1}$$

Where D is the number of all documents in the collection while df_j is the number of documents containing the term TF_{ij}. Thus the general formula for term weighting TF × IDF is a combination of the TF formula with the IDF formula by multiplying the TF value by the IDF value as in **equation 2** [11]. Where w_{ij} is the weight of the term against the documents. Regardless of the value of TF_{ij}, if D is equal to df_j, then 0 (zero) will be obtained because the result is log 1. For this reason, a value of 1 can be added on the IDF side.

$$w_{ij} = TF_{ij}\; x \log\left(\frac{D}{df_j}\right) + 1 \tag{2}$$

Vector word can be interpreted as a word vector from a process of making existing sentences into a collection of arrays that are collected into a matrix, where each row of the matrix represents a row of the document, and each column of the matrix represents all the words in the text of the data. After all the words are processed and become vectors, the word weighting process is carried out using the TF-IDF method.

TABLE 2. TF-IDF calculation

	TF			DF	D/DF	IDF+1	W = TF * (IDF+1)		
	Doc1	Doc2	Doc3				Doc1	Doc2	Doc3
Belajar	1	1	1	3	1	1	1	1	1
Daring	1	1	1	3	1	1	1	1	1
Stress	0	1	0	1	3	1.477	0	1.477	0
Semangat	0	0	1	1	3	1.477	0	0	1.477
The weight value of each document							2	3.477	3.477

This study uses the Naive Bayes Classifier method to classify data to get the sentiment analysis results. Naive Bayes has proven effective in many practical applications, including text classification, medical diagnosis, and systems performance management [12]. Required data that has been preprocessed and word-weighted will then be used as training data before proceeding to the testing stage to get the classification results [13]. For calculation, the probability of document C to classify on a class **equation 3** will be used to obtain the probability score of document C.

$$P(C|F_1, \dots, F_n) = P(C)\, P(F_1|C) P(F_2|C)\, P(F_3|C) \dots\, P(F_n|C) \tag{3}$$

For calculation, the probability of category C **equation 4** will be used to obtain the probability score in some categories.

$$P(c) = \frac{N_c}{N} \tag{4}$$

For calculation, the probability of a word from category C **equation 5** will be used to obtain the probability score from the word in some category.

$$P(F_n|C) = \frac{count(tn,c)+1}{count\ (c)+|V|} \tag{5}$$

RESULTS AND DISCUSSION

The model test was carried out with several scenarios. The first scenario divides the training data by 80% and the test data by 20%. The results obtained are shown in **TABLE 3**.

The values results of accuracy, precision, recall, and F1 score have an assessment size between 0-1. The higher the resulting value, the better. This means that the closer the value is to 1, the better the model used has a good performance on the data used. The results of the values of accuracy, precision, recall, and F1 score in the first scenario can be seen in **TABLE 4**.

TABLE 3. Scenario 1 Confusion Matrix Results

Actual Class	Predicted Class	
	Positive	Negative
Positive	84	19
Negative	42	55

TABLE 4. Scenario 1 Classification Results

Scenario	Accuracy	Precision	Recall	F1 Score
1	0.695	0.743	0.567	0.643

The second scenario with 70% training data and 30% test data. The results of the confusion matrix are obtained, which are shown in **TABLE 5**. Then the classification results for the second scenario can be seen in **TABLE 6**.

TABLE 5. Scenario 2 Confusion Matrix Results

Actual Class	Predicted Class	
	Positive	Negative
Positive	132	23
Negative	73	72

TABLE 6. Scenario 2 Classification Results

Scenario	Accuracy	Precision	Recall	F1 Score
2	0.680	0.757	0.496	0.600

The third scenario with 60% training data and 40% test data results in the confusion matrix, as shown in **TABLE 7**, then the classification results can be seen in **TABLE 8**.

TABLE 7. First Confusion Matrix Scenario Results

Actual Class	Predicted Class	
	Positive	Negative
Positive	170	51
Negative	83	96

TABLE 8. Scenario 1 Classification Results

Scenario	Accuracy	Precision	Recall	F1 Score
3	0.665	0.653	0.536	0.588

Based on the calculation results from the first to the third scenario, the classification evaluation was obtained from the values of accuracy, precision, recall, and F1 score, summarized in **FIGURE 3**.

As shown in **FIGURE 3**, the highest accuracy value was obtained during Scenario 1, which was 0.695 with 80% training data and 20% test data. Then the precision value is 0.743. The recall value is 0.567. The F1 score is 0.643

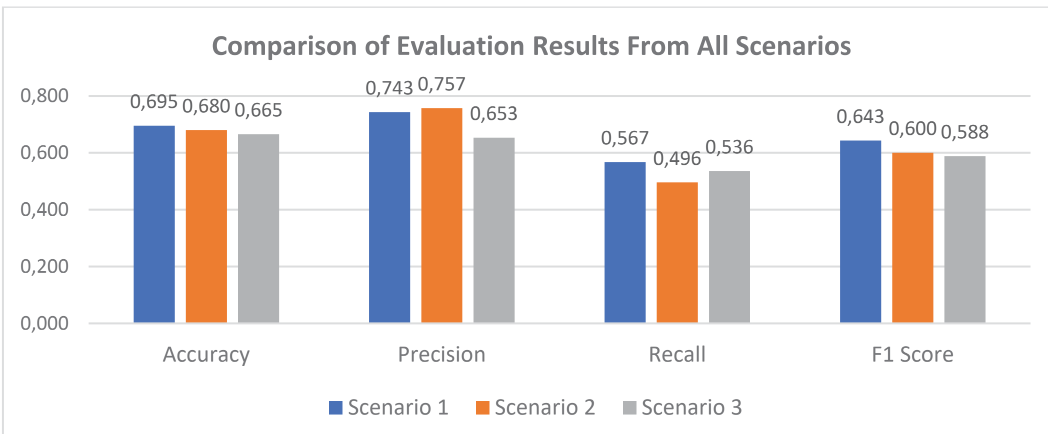

FIGURE 3. Comparison of Evaluation Results

At this calculation stage, the existence of a stemming process can also affect the classification results, where a comparison is also made between the classification process with the stemming process and the classification process without the stemming process. The comparison data is taken from the highest accuracy value from the previous classification process, Scenario 1, with training data of 80% and test data of 20%, with an accuracy value of 0.695.

As shown in **FIGURE 4,** the comparison is taken from scenario 1. Recalculation is carried out to compare and prove whether the stemming process affects the classification results.

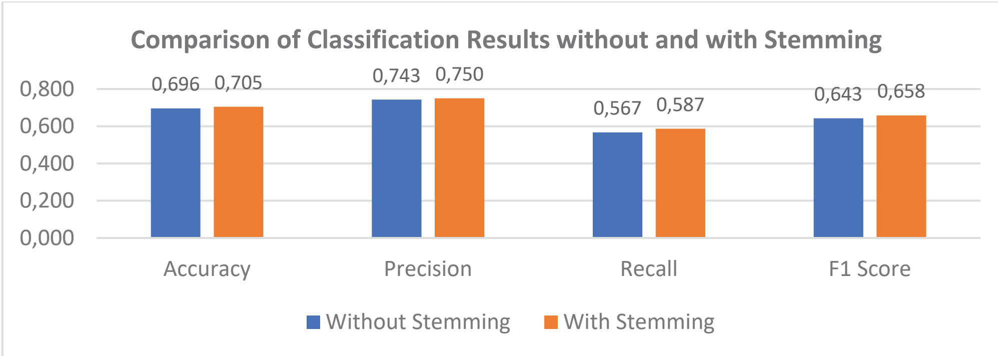

FIGURE 4. Comparison of Classification Results

The results obtained are that through the stemming process, the classification results from sentiment analysis have slightly increased in terms of accuracy, precision, recall, and F1 scores. It can be seen that the accuracy value has increased with a value of 0.705, precision of 0.750, recall of 0.587, and F1 score of 0.658. So it can be said that the stemming process can also affect the final value of the classification results.

CONCLUSIONS

In this study, a system has been developed to analyze the sentiment of the impact of COVID-19 on education in Indonesia using Twitter social media data. The Naive Bayes Classifier algorithm has an accuracy rate of 0.696 with a process without the stemming stage and 0.705 through the stemming stage. The presence or absence of the stemming process is also quite influential on the final result of the classification. In this study, the final value of the classification slightly increased when going through the stemming process compared to without going through the stemming stage. Further research can be conducted to investigate the semantic relationship in tweets where words can be biased or have multiple meanings in a tweet.

REFERENCES

[1] R. H. S. Aji, "Dampak COVID-19 pada pendidikan di indonesia: Sekolah, keterampilan, dan proses pembelajaran," *Jurnal Sosial & Budaya Syar-i,* vol. 7, no. 5, pp. 395-402, 2020.

[2] R. Pangaribuan, "Dampak Positif Pandemi Covid-19 Terhadap Dunia Pendidikan," *STIT Al-Kifayah Riau, Mar,* vol. 8, 2021.

[3] G. A. Buntoro, "Analisis Sentimen Calon Gubernur DKI Jakarta 2017 Di Twitter," *INTEGER: Journal of Information Technology,* vol. 2, no. 1, 2017.

[4] G. N. Aulia and E. Patriya, "Implementasi Lexicon Based Dan Naive Bayes Pada Analisis Sentimen Pengguna Twitter Topik Pemilihan Presiden 2019," *Jurnal Ilmiah Informatika Komputer,* vol. 24, no. 2, pp. 140-153, 2020.

[5] A. Rahmatulloh, N. I. Kurniati, I. Darmawan, A. Z. Asyikin, and D. Witarsyah, "Comparison between the stemmer porter effect and nazief-adriani on the performance of winnowing algorithms for measuring plagiarism," *International Journal on Advanced Science, Engineering and Information Technology,* vol. 9, no. 4, pp. 1124-1128, 2019.

[6] M. Adriani, J. Asian, B. Nazief, S. M. M. Tahaghoghi, and H. E. Williams, "Stemming Indonesian: A confix-stripping approach," *ACM Transactions on Asian Language Information Processing (TALIP),* vol. 6, no. 4, pp. 1-33, 2007.

[7] I. Prismana, D. R. Prehanto, D. A. Dermawan, A. C. Herlingga, and S. C. Wibawa, "Nazief & Adriani Stemming Algorithm With Cosine Similarity Method For Integrated Telegram Chatbots With Service," vol. 1125, p. 012039: IOP Publishing.

[8] A. Aizawa, "An information-theoretic perspective of tf–idf measures," *Information Processing & Management,* vol. 39, no. 1, pp. 45-65, 2003.

[9] S. Qaiser and R. Ali, "Text mining: use of TF-IDF to examine the relevance of words to documents," *International Journal of Computer Applications,* vol. 181, no. 1, pp. 25-29, 2018.

[10] C.-z. Liu, Y.-x. Sheng, Z.-q. Wei, and Y.-Q. Yang, "Research of text classification based on improved TF-IDF algorithm," pp. 218-222: IEEE.

[11] I. Yahav, O. Shehory, and D. Schwartz, "Comments mining with TF-IDF: the inherent bias and its removal," *IEEE Transactions on Knowledge and Data Engineering,* vol. 31, no. 3, pp. 437-450, 2018.

[12] D. Berrar, "Bayes' theorem and naive Bayes classifier," *Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics,* vol. 403, 2018.

[13] F.-J. Yang, "An implementation of naive bayes classifier," pp. 301-306: IEEE.