International Applied Science
**UMJember Proceeding Series** (2023) Vol. 2, No. 1 : 14-20

**SIGMA-1 EKSAKTA**

# *Sentiment Analysis Tweets* Against the Independent Learning Curriculum Policy on *Twitter Social Media* Using the *Mutivariate Bernoulli Algorithm*

**Tedi Erwanto [1] , Bagus Setya Rintyarna [2] and Nur Qodariyah Fitriyah [3]**
[1] Muhammadiyah University of Jember ; tedierwantosm@gmail.com
[2] Muhammadiyah University of Jember ; baik.setya@bagus.setya.ac.id
[3] Universitas Muhammadiyah Jember; nurfitriyah@unmuhjember.ac.id

*Correspondensi: Tedi Erwanto
Email: tedierwantosm@gmail.com

Published: January, 2023



**Abstract:** Merdeka Belajar is a new policy initiated by the Ministry of Education, Culture, Research and Technology. According to him, this policy is a step to change the education system in Indonesia so that Human Resources (HR) in Indonesia become superior and have a Pancasila Student profile . These policy efforts cannot be separated from support and rejection from the community. Various kinds of support and rejection efforts from the public regarding the new policy were expressed on social media, one of which was Twitter . This research aims to determine the level of *accuracy, precision* and *recall* using the *Multivariate Bernoulli* algorithm with extra TF -IDF features. The highest *accuracy* result was 81.83%, in follow *precision* of 74.24%, and *recall* of 84.48%. Average performance on 10- Fold Cross *Validation* provided by *Confusion Matrix* with *accuracy values* of 78.00%, *precision* 73.73%, and *recall 81.09%, as well as obtaining* 1,092 correctly classified data results and 308 data were classified incorrectly

**Keywords:** Twitter, Merdeka Belajar, Sentiment Analysis, TF-IDF, *Multivariate Bernoulli*

## INTRODUCTION

The Ministry of Education, Culture, Research and Technology launched a new policy, namely Merdeka Belajar, according to him, this policy is a step to change the education system in Indonesia so that human resources (HR) in Indonesia are superior and have a Pancasila Student profile. This policy cannot be separated from support and rejection from the community. This support and rejection is often called sentiment.

*Sentiment analysis* is a process of analyzing a person's opinions, emotions and feelings expressed through text and processing data from unstructured to structured data. In its application, the sentiment analysis process will produce negative, positive and neutral output (Nugroho et al., 2021). To find out whether the comments fall into positive, negative or neutral comments, a classification process will be applied.

Various methods applied in sentiment analysis classification have different accuracy values. From the classification process, *accuracy results will be obtained* from processing using the method applied. Several types of methods that are often applied, one of which is *Multivariate Benoulli* , where this method calculates each test data with the probability value of each word in each class. Calculation of probability in the *Multivariate Bernoull* i model using data containing words rather than the frequency of word occurrence (Karunia et al., 2017).

The aim of this research is to determine the results *of accuracy, precision* and *recall* using the *Multivariate Bernoulli method* with TF-IDF feature extraction. TF-IDF is the process of changing words into numerical form. This process is used to measure how far the words in a document are related by giving a value to each word. TF-IDF is a combination of two concepts, namely TF is the number of words that appear in a document and IDF is the frequency of words that appear in a document (Deolika et al., 2019).

Previous research conducted by Widyawati and Susanto in 2020 discussed the performance comparison between the *Multivariate Bernoulli algorithm* and *Multinomial Naïve Bayes* , the scenario applied was the influence of the use of *text preprocessing* ( *stemming* ) on the accuracy results and the time required in the classification process. The results of this research were that the highest *accuracy* was obtained by applying the *Bernoulli Naïve Bayes non stemming method* of 71.33% and the fastest time in the classification process was applying the *Multinomial Naïve Bayes non stemming method* , namely with a processing time of 0.12 seconds. Then research conducted by Karunia in 2017 discussed online news classification using *the Naïve Bayes Classifier with Mutual Information* feature selection . The best results from this method are *accuracy* 80%, *precision* 94.28%, *recall* 79.68% and *f-measure* 85.08% for *Multivariate Bernoulli* without feature selection. Then the best results from the classification model with feature selection were achieved in the *Multivariate Bernoulli* model with *an accuracy of* 70%, *precision* 89.11%, *recall* 69.76% and *f-measure* 78.04% with a word efficiency level of up to 52% than before, using selection feature. Meanwhile, the results of *Multinomial Naïve Bayes* without feature selection are *accuracy* 41.67%, *precision* 75.68%, *recall* 41.90% and *f-measure* 48.13%, for the results of the *Multinomial Naïve Bayes model* with feature selection 10% *accuracy* , 33, 33% *precision* , 9.40% *recall* and 14.35% *f-measure* . However, in this research the data *preprocessing techniques* used were only *cleansing, case folding, stopword removal, stemming,* and *tokenizing,* where in a sentence there were writing errors ( *typos* ) and non-standard words which could not be overcome by some of these *preprocessing text techniques* . Therefore, the novelty presented in this research applies the Normalization *text preprocessing technique* to change writing errors ( *typos* ) and non-standard words, with the hope of increasing *accuracy results* that are higher compared to previous research.

## METHOD

Research on *tweet sentiment analysis* Regarding the independent learning curriculum policy using the *Multivariate B Bernoulli algorithm* , there are several important steps that must be taken, namely starting from collecting *tweet data* using crawling techniques, then labeling the data, then the data is processed at the *preprocessing stage* , word weighting using TF-IDF feature extraction, then the stage of validation and implementation of *the Multivariate Bernoulli algorithm* . In general, an overview of the steps in this research can be seen in Figure 1 below :

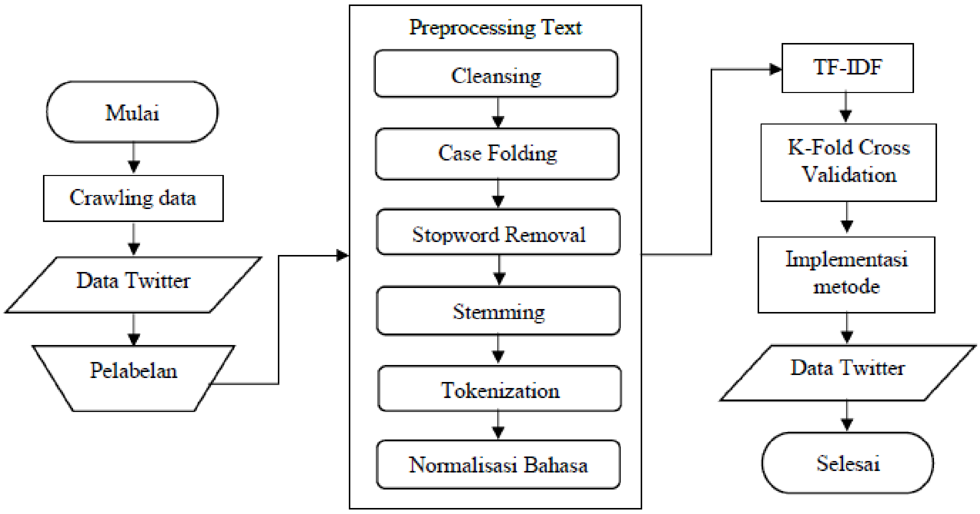


**Figure 1.** Research Stages

### Data collection

Data collection was carried out using the *crawling technique. Crawling* is a data collection process carried out by retrieving data via Twitter social media with data crawling techniques using the Twitter API ( *Application Programming Interface* ). The data retrieval process uses the Python programming language. The data collection process uses the keyword "independent learning". The data collected was 1400 data taken in June - August 2022. Labeling was done manually and validated by a linguist.

### *Text Preprocessing*

*Preprocessing Text* is an important step to prepare data in the form of text so that it is ready for processing at the next stage. The stages that will be carried out are as follows:

a. Cleansing

This process will delete words that are unimportant or have no meaning in a sentence, such as punctuation, symbols, hashtags, URLs, etc. With the aim of minimizing the occurrence of noise during processing (Saraswati & Riminarsih, 2020) .

b. Case Folding
This process will change uppercase letters to lowercase in a sentence so that the text to be processed has the same letters (Haditira et al., 2022) .

c. Stopword Removal
This process will remove conjunctions contained in a sentence which are considered to have no meaning or meaning in a sentence such as: or, in, to, and, etc. (Haditira et al., 2022) .

d. Stemming
This process will involve changing words that add at the beginning of a word or at the end of a word as well as at the beginning and end of a word so that it becomes a basic word such as: "serve" becomes "servant", "feel" becomes "rasa", "calculate" becomes "count", and so on (Saraswati & Riminarsih, 2020) .

e. Normalization
This process will change non-standard words into standard words, change abbreviations into proper words, change words with misspellings or typos, change slang or what is usually called a word that is not included in the Big Indonesian Dictionary (KBBI) (Ivan & Adikara, 2019) .

f. Tokenizing
This process will separate each word in a sentence so that it becomes a word arrangement (Nurhuda et al., 2016) .

### *Term Frequency-Inverse Document Frequency )* Feature Extraction

*Term Frequency-Inverse Document Frequency* (TF-IDF) is a technique for weighting each word ( *term* ) in a document that has undergone preprocessing.

$$\mathrm{W}tf_{t,d} = \begin{cases} 1 + log_{10} tf_{t,d}, if\ tf_{t,d} > 0 \\ 0, \qquad\qquad\quad if\ tf_{t,d} = 0 \end{cases} \tag{1}$$

$$idf_t = log_{10} \frac{N}{df_t} \tag{2}$$

$$W_{t,d} = Wtf_{t,d} \,.\, idf_t \tag{3}$$

### *K-Fold Cross Validation*

*K-Fold Cross Validation* process is by dividing the data set into several parts consisting of *training data* and *testing data* . In this study, 4 types were used *K-Fold* .

1. 2- *Fold Cross Validation*
   2- *Fold Cross Validation* is data divided into 2 equal parts, each part has a total of 700 data.

$$\frac{banyak\ data}{k\ fold} = \frac{1400}{2} = 700$$

So 700 data will be used as *testing data* and the rest will be used as *training data* .

2. 5- *Fold Cross Validation*
   5- *Fold Cross Validation* is data divided into 5 equal parts, each part has a total of 280 data.

$$\frac{banyak\ data}{k\ fold} = \frac{1400}{5} = 280$$

So 280 data will be used as *testing data* and the rest will be used as *training data* .

3. 7- *Fold Cross Validation*

7- *Fold Cross Validation* is data divided into 7 equal parts so that each part has a total of 200 data.

$$\frac{banyak\ data}{k\ fold} = \frac{1400}{7} = 200$$

So 200 data will be used as *testing data* and the rest will be used as *training data* .

*4.* 10- *Fold Cross Validation*

10- *Fold Cross Validation* is data will divided into 10 equal parts so that each part has a total of 140 data.

$$\frac{banyak\ data}{k\ fold} = \frac{1400}{10} = 140$$

So 140 data will be used as *testing data* and the rest will be used as *training data* .

***Multivariate Bernoulli* Implementation**

*Multivariate Bernoulli* calculates each test data with the probability value of each word in each class. Calculation of probability in the *Multivariate Bernoulli model* using data containing the word t (Tct) rather than the frequency of occurrence of the word t . *Bernoulli* model using the amount of training data for each class (Tc) and the number of classes ∑c. According to (Karunia et al., 2017) the probability of each word is calculated using the equation:

$$P(fk|c) = \frac{Tct+1}{Tc+\sum c} \tag{4}$$

After the probability of each word is obtained, the value of the inverse product for each probability of the word in the training data other than the word in the test data is calculated. Next, look for words in the training data d that are independent of the words in the test data f. From these words, the probability value of the word is then taken (in this case given the symbol P(fk'|c)) and the number of words (M). The probability value of a sentence for a class is calculated using the equation

$$P(c|d) = P(c)\ \prod_{i=1}^{N} P(fk|c)\ x \prod_{i=1}^{M}(1 - P(fk'|c)) \tag{5}$$

## RESULTS AND DISCUSSION

The following are the *accuracy*, *precision* and *recall values* obtained in *fold cross validation* for each *fold* . The classification technique used in all tests uses the *Multivariate Bernoulli algorithm.*

**Table 1.** *Accuracy, precision and recall* results

| *K-Fold Cross* | Test Steps | *accuracy* | *precision* | *recalls* |
|---|---|---|---|---|
| 2- *Fold* | Test Step 1 | 76.14% | 74.21% | 73.52% |
| | Test Step 2 | 75.86% | 72.67% | 75.62% |
| | Average | 76.00% | 73.44% | 74.57% |
| 5- *Fold* | Test Step 1 | 79.29% | 76.81% | 80.30% |
| | Test Step 2 | 78.93% | 73.61% | 83.46% |
| | Test Step 3 | 74.64% | 69.23% | 78.57% |
| | Test Step 4 | 77.50% | 71.43% | 79.17% |
| | Test Step 5 | 79.29% | 77.86% | 80.15% |
| | Average | 77.93% | 73.79% | 80.33% |
| 7- *Fold* | Test Step 1 | 81.00% | 77.45% | 84.04% |
| | Test Step 2 | 79.00% | 75.96% | 82.29% |
| | Test Step 3 | 79.50% | 74.49% | 82.02% |
| | Test Step 4 | 72.00% | 64.29% | 81.82% |
| | Test Step 5 | 78.00% | 75.28% | 75.28% |
| | Test Step 6 | 75.50% | 67.92% | 82.76% |

| *K-Fold Cross* | Test Steps | *accuracy* | *precision* | *recalls* |
|---|---|---|---|---|
| | Test Step 7 | 79.50% | 78.79% | 79.59% |
| | Average | 77.79% | 73.45% | 81.12% |
| 10- *Fold* | Test Step 1 | 81.43% | 74.24% | 84.48% |
| | Test Step 2 | 80.71% | 82.19% | 81.08% |
| | Test Step 3 | 77.14% | 72.37% | 83.33% |
| | Test Step 4 | 79.29% | 72.86% | 83.61% |
| | Test Step 5 | 74.29% | 67.57% | 80.65% |
| | Test Step 6 | 77.86% | 74.63% | 78.12% |
| | Test Step 7 | 77.14% | 68.57% | 82.76% |
| | Test Step 8 | 77.86% | 73.13% | 79.03% |
| | Test Step 9 | 75.00% | 68.42% | 82.54% |
| | Test Step 10 | 79.29% | 83.33% | 75.34% |
| | Average | 78.00% | 73.73% | 81.09% |

Based on the results recap *accuracy* , *precision* and recall in Table 1, the highest *accuracy* with a value of 81.43% was obtained on the 10th *Fold in test Step 1. On the* 10th *Fold in Test Step 1 , the precision value was* 74.24%, and *the recall* was 84.48 %. Then the lowest *accuracy* with a value of 72.00% was obtained on the 7th *Fold in test Step 4. On* the 7th *Fold in test Step 4, the precision value was* 64.29%, and *the recall* was 81.82%. Average performance on the 10th *Fold Cross Validation with accuracy values of* 78.00%, *precision* 73.73%, and *recall* 81.09%.

Evaluation of the performance of the classification process is measured using *the Confusion Matrix* to see the accuracy of *Multivariate Bernoulli* in classifying data in each class to be classified. The following is one of *the Confusion Matrix* on the 10th *Fold. Test step 1 can be seen in Table 2.*

**Table 2.** *Confusion Matrix Fold* -10 Test steps 1

| | Actual Positive (+) | Actual Negative (-) | **Amount** |
|---|---|---|---|
| Prediction (+) | 49 | 9 | **58** |
| Prediction (-) | 17 | 65 | **82** |
| **Amount** | **66** | **74** | **140** |

In Table 2 it can be seen that *Multivariate Bernoulli* on *Fold* -10 Test step 1 there are 66 data that are predicted to be positive and 74 data that are predicted to be negative. However, of the 66 data, there was a prediction error of 17 data, of which 17 data should be in the positive class. For the 74 data that were predicted to be negative, there were 9 data that were predicted incorrectly, where the data should have been in the negative class. Therefore, the number of data that was classified correctly was 114 and 26 data were classified incorrectly.

Overall results of the *K-Fold Cross Validation experiment* using the *Multivariate algorithm Bernoulli* can say that the comparison of the amount of training data with test data affects *the accuracy* , *precision* and *recall results* . The more training data used, the higher the *accuracy results* obtained. In this research, the amount of training data on 10- *Fold* was 1,260 and the test data was 140 data. The *accuracy* results obtained increased by 1.43% compared to previous research.

## CONCLUSION

The aim of the research is to determine *the accuracy, precision* and *recall* of the independent learning curriculum policy using the *Multivariate Bernoulli algorithm.* The results of experimentation and validation using the *K-Fold Cross Validation technique* by adding *text preprocessing* Normalization obtained the highest *accuracy* with a value of 81.43%, followed by a *precision value* of 74.24%, and *a recall* of 84.48%.

Average performance on the 10th fold cross validation provided by *Confusion Matrix* with *accuracy values of* 78.00%, *precision* 73.73%, and *recall* 81.09%, as well as obtaining correctly classified data results of 1,092 data and 308 data was classified incorrectly in 10- *fold cross validation* . Based on these results, the *Multivariate Bernoulli Algorithm* can be used for *sentiment analysis* of the independent learning curriculum policy, and has good performance in carrying out the classification process.

## REFRENCES

Al Kabir, A. H., Basuki, S., & Wicaksono, G. W. (2019). Analisis sentimen kritik dan saran pelatihan aplikasi teknologi informasi (PATI) menggunakan algoritma support vector machine (SVM). Jurnal Repositor, 1(1), 39. https://doi.org/10.22219/repositor.v1i1.11

Anggreini, N. M., Nasir, B. M. S., & Noor, I. L. S. Sos. M. (2016). Pemanfaatan Media Sosial Twitter di Kalangan Pelajar SMK Negeri 5 Samarinda. Sosiologi, 4(2), 243.

Arisandi, R. R. R., Warsito, B., & Hakim, A. R. (2022). Aplikasi Naïve Bayes Classifier (Nbc) Pada Klasifikasi Status Gizi Balita Stunting Dengan Pengujian K-Fold Cross Validation. Jurnal Gaussian, 11(1), 130–139. https://doi.org/10.14710/j.gauss.v11i1.33991

Ashari, H., Arifianto, D., Azizah, H., & Faruq, A. (n.d.). PERBANDINGAN KINERJA ALGORITMA MULTINOMIAL NAÏVE BAYES (MNB), MULTIVARIATE BERNOULLI DAN ROCCHIO ALGORITHM DALAM KLASIFIKASI KONTEN BERITA HOAX BERBAHASA INDONESIA PADA MEDIA SOSIAL.

Azizah, H., Rintyarna, B. S., & Cahyanto, T. A. (2022). Sentimen Analisis Untuk Mengukur Kepercayaan Masyarakat Terhadap Pengadaan Vaksin Covid-19 Berbasis Bernoulli Naive Bayes. BIOS: Jurnal Teknologi Informasi Dan Rekayasa Komputer, 3(1), 23–29. https://doi.org/10.37148/bios.v3i1.36

Deolika, A., Kusrini, K., & Luthfi, E. T. (2019). Analisis Pembobotan Kata Pada Klasifikasi Text Mining. Jurnal Teknologi Informasi, 3(2), 179. https://doi.org/10.36294/jurti.v3i2.1077

Giovani, A. P., Ardiansyah, A., Haryanti, T., Kurniawati, L., & Gata, W. (2020). Analisis Sentimen Aplikasi Ruang Guru Di Twitter Menggunakan Algoritma Klasifikasi. Jurnal Teknoinfo, 14(2), 115. https://doi.org/10.33365/jti.v14i2.679

Haditira, R., Informatika, F., Telkom, U., Review, S., Murdiansyah, D. T., Informatika, F., Telkom, U., Astuti, W., Informatika, F., & Telkom, U. (2022). Analisis Sentimen Pada Steam Review Menggunakan Metode Multinomial Naïve Bayes dengan Seleksi Fitur Gini Index Text. 9(3), 1793–1799.

Harismawan, A. F. (2017). Analisis Perbandingan Performa Web Service Menggunakan Bahasa Pemrograman Python, Php ,. Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer, 2(1), 237–245.

Husada, H. C., & Paramita, A. S. (2021). Analisis Sentimen Pada Maskapai Penerbangan di Platform Twitter Menggunakan Algoritma Support Vector Machine (SVM). Teknika, 10(1), 18–26. https://doi.org/10.34148/teknika.v10i1.311

Ivan, Y. A. S., & Adikara, P. P. (2019). Klasifikasi Hate Speech Berbahasa Indonesia di Twitter Menggunakan Naive Bayes dan Seleksi Fitur Information Gain dengan Normalisasi Kata. Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer, 3(5), 4914–4922.

Karunia, S. A., Saptono, R., & Anggrainingsih, R. (2017). Online News Classification Using Naive Bayes Classifier with Mutual Information for Feature Selection. Jurnal Ilmiah Teknologi Dan Informasi, 6(1), 10–15.

Makhmudah, U. (2019). Analisis Sentimen Terhadap Tweet Kaum Homoseksual Indonesia Menggunakan Metode Support Vector Machine. In Repository.Unej.Ac.Id.

Nugroho, D. A., Savitri, N., Saputra, N., Adji, T. B., Permanasari, A. E., Ardhiansyah, M. N., Umar, R., Sunardi, Alifah, U., Arsya Monica Pravina, Imam Cholissodin, P. P. A., Darwis, D., Pratiwi, E. S., Pasaribu, A. F. O., Arsi, P., & Waluyo, R. (2021). Analisis Sentimen Data Presiden Jokowi Dengan Preprocessing Normalisasi Dan Stemming Menggunakan Metode Naive Bayes Dan SVM. Nasional Teknologi Fakultas Teknik Universitas Krisnadwipayana, 3(1), 1–11.

Nurhuda, F., Widya Sihwi, S., & Doewes, A. (2016). Analisis Sentimen Masyarakat terhadap Calon Presiden Indonesia 2014 berdasarkan Opini dari Twitter Menggunakan Metode Naive Bayes Classifier. Jurnal Teknologi & Informasi ITSmart, 2(2), 35. https://doi.org/10.20961/its.v2i2.630

Pintoko, B. M., & L., K. M. (2018). Analisis Sentimen Jasa Transportasi Online pada Twitter Menggunakan Metode Naive Bayes Classifier. E-Proceeding of Engineering, 5(3), 8121–8130.

Prasetyo, H. D., Pramiyati, T., & Isnainiyah, I. N. (2021). Sentimen Analisis Pengguna Twitter Terhadap Kebijakan Merdeka Belajar Menggunakan Algoritma Naive Bayes. Seminar Nasional Mahasiswa Ilmu Komputer Dan Aplikasinya (SENAMIKA), April, 559–568.

Rozaq, A., Yunitasari, Y., Sussolaikah, K., Resty, E., Sari, N., & Syahputra, R. I. (2022). Analisis Sentimen Terhadap Implementasi Program Merdeka Belajar Kampus Merdeka Menggunakan Naïve Bayes, K-Nearest Neighboars Dan Decision Tree. 6(April), 746–750. https://doi.org/10.30865/mib.v6i2.3554

Saraswati, M., & Riminarsih, D. (2020). Analisis Sentimen Terhadap Pelayanan Krl Commuterline Berdasarkan Data Twitter Menggunakan Algortima Bernoulli Naive Bayes. Jurnal Ilmiah Informatika Komputer, 25(3), 225–238. https://doi.org/10.35760/ik.2020.v25i3.3256

Sari, B. W., & Haranto, F. F. (2019). IMPLEMENTASI SUPPORT VECTOR MACHINE UNTUK ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP PELAYANAN TELKOM DAN BIZNET. Jurnal Pilar Nusa Mandiri, 15(2), 171–176. https://doi.org/10.33480/pilar.v15i2.699

Widyawati, & Sutanto. (n.d.). PERBANDINGAN KINERJA VARIASI NAÏVE BAYES MULTIVARIATE BERNOULLI DAN NAÏVE BAYES MULTINOMIAL DALAM PENGKLASIFIKASIAN DOKUMEN TEKS.