

**Implementasi Algoritma *Multinomial Naïve Bayes* Untuk
Analisis Sentimen Twitter Terhadap Kebijakan Merdeka Belajar
*Implementation Of Naïve Bayes Multinomial Algorithm
For Sentiment Analysis Of Independent Learning Policies***

Aswa Salsabilla Wilindia¹, Moh. Dasuki², Nur Qodariyah Fitriyah³

¹ Masiswa Program Studi Teknik Informatika, Fakultas Teknik, Universitas Muhammadiyah Jember

Email : aswasalsabilla3@gmail.com

² Dosen Fakultas Teknik, Universitas Muhammadiyah Jember

Email : moh.dasuki22@unmuhjember.ac.id

³ Dosen Fakultas Teknik, Universitas Muhammadiyah Jember

Email : nurfitriyah@unmuhjember.ac.id

Abstrak

Merdeka belajar adalah program yang dibuat oleh Kementerian Pendidikan dan Kebudayaan Republik Indonesia (Kemendikbud RI) yang diumumkan pada tahun 2019 oleh Menteri Pendidikan dan Kebudayaan Republik Indonesia Kabinet Indonesia Maju, Nadiem Anwar Makarim. Kebijakan merdeka belajar bertujuan untuk memberikan kebebasan kepada siswa dalam proses pembelajaran, untuk meningkatkan kemampuannya dalam memenuhi kebutuhan zaman yang semakin modern seiring dengan perkembangan teknologi yang signifikan. Namun, pencapaian program merdeka belajar menimbulkan banyak komentar pro dan kontra dari masyarakat luas, yang disampaikan secara langsung maupun melalui media sosial. Twitter merupakan salah satu media sosial yang banyak digunakan untuk menyampaikan opini dan aspirasi dari masyarakat luas menggunakan pesan yang biasa disebut dengan *tweet*, yang bersifat positif, negatif dan netral. Tujuan dari penelitian ini adalah untuk melakukan analisis sentimen opini publik terhadap kebijakan merdeka belajar, serta untuk mengetahui hasil akurasi, presisi dan *recall* yang paling optimal menggunakan algoritma *Multinomial Naïve Bayes* dengan ekstraksi fitur *TF-IDF*. Hasil penelitian menunjukkan nilai akurasi tertinggi pada *fold* $k=10$ pada langkah uji 5 dengan hasil 75.38%, presisi 76.32%, dan *recall* 80.56%.

Kata Kunci : Merdeka Belajar, Twitter, Analisis Sentimen, *Multinomial Naïve Bayes*, *TF-IDF*

Abstract

Freedom to learn is a program created by the Ministry of Education and Culture of the Republic of Indonesia (Kemendikbud RI) which was announced in 2019 by the Minister of Education and Culture of the Republic of Indonesia Advanced Indonesian Cabinet, Nadiem Anwar Makarim. The independent learning policy aims to give freedom to students in the learning process, to improve their ability to meet the needs of an increasingly modern era along with significant technological developments. However, the achievement of the independent learning program has generated many pro and con comments from the wider community, which were conveyed directly or through social media. Twitter is one of the social media that is widely used to convey opinions and aspirations from the wider community using messages commonly called tweets, which are positive, negative and neutral. The purpose of this research is to carry out an analysis of public opinion sentiment towards the independent learning policy, and to find out the results of the most optimal accuracy, precision and recall using the Multinomial Naïve Bayes algorithm with TF-IDF feature extraction. The results showed the highest accuracy value at fold $k = 10$ in test step 5 with a yield of 75.38%, 76.32% precision, and 80.56% recall.

Keywords : *Freedom to Learn, Twitter, Sentiment Analysis, Multinomial Naïve Bayes, TF-IDF*

1. PENDAHULUAN

Kebijakan merdeka belajar adalah program yang dibuat oleh Kementerian Pendidikan dan Kebudayaan Republik Indonesia (Kemendikbud RI) yang diumumkan pada tahun 2019 oleh Menteri Pendidikan dan Kebudayaan Republik Indonesia Kabinet Indonesia Maju, Nadiem Anwar Makarim. Program merdeka belajar dapat memungkinkan siswa untuk lebih mengembangkan keterampilan sesuai dengan bakat dan minat mereka. Kebijakan merdeka belajar bertujuan untuk memberikan kebebasan kepada siswa dalam proses pembelajaran.

Pencapaian program merdeka belajar ini tentunya menimbulkan pro dan kontra dari masyarakat luas. Para akademisi maupun masyarakat umum mengkritik adanya kebijakan merdeka belajar, baik secara langsung maupun melalui media sosial. Twitter adalah salah satu *platform* jejaring sosial yang banyak digunakan untuk membentuk opini di media sosial. Twitter adalah layanan *microblogging* dan jejaring sosial yang memungkinkan pengguna untuk memposting dan berinteraksi dengan pesan yang disebut “*tweet*”.

Analisis sentimen adalah proses menemukan teks yang diungkapkan oleh pengguna twitter. Dalam hal ini, teks dapat memiliki makna positif, negatif, dan bahkan netral. Banyak metode klasifikasi yang digunakan untuk analisis sentimen, seperti *Multinomial Naïve Bayes*. *Multinomial Naïve Bayes* adalah pengembangan dari algoritma *bayesian* yang biasa digunakan dalam klasifikasi teks. Sebagaimana penelitian (Haditira et al., 2022) yang berjudul “Analisis Sentimen Pada *Steam Review* Menggunakan Metode *Multinomial Naïve Bayes* dengan Seleksi Fitur *Gini Index Text*” tujuan dari penelitian tersebut adalah memproses data serta melakukan seleksi fitur pada sentimen analisis *steam review* dan mengukur performa sistem. Dengan hasil rata – rata akurasi diperoleh dengan metode *Multinomial Naïve Bayes* tanpa seleksi fitur *Gini Index Text* adalah 81.38%, 83.30%, 80.08%, dan 81.04%. Rata – rata akurasi dengan seleksi fitur *Gini Index Text* adalah 60.29%, 56.65%, 91.46%, 69.71% dengan nilai maksimal 60.29%. Penelitian selanjutnya (Prasetyo et al., 2021) dengan judul penelitian “Sentimen Analisis

Pengguna Twitter Terhadap Kebijakan Merdeka Belajar Menggunakan Algoritma *Naïve Bayes*” dengan hasil penelitian 180 data “*tweet*” dibagi menjadi data latih dan data uji, dengan nilai akurasi terbesar 80.55% untuk nilai *f1-score* 89% nilai presisi 81%, dan *recall* 100% dengan data latih sebesar 80% dan data uji sebesar 20%.

Berdasarkan latar belakang di atas, penelitian ini melakukan analisis tentang opini publik terhadap kebijakan merdeka belajar berbasis data twitter. Dengan menerapkan metode klasifikasi *Multinomial Naïve Bayes* dengan ekstraksi fitur *TF-IDF*. Dengan menggunakan *k-fold cross validation* dengan $k=2, 5, \text{ dan } 10$ agar mendapatkan perbandingan yang adil antara data latih dan data uji. Dengan judul penelitian “Implementasi Algoritma *Multinomial Naïve Bayes* untuk Analisis Sentimen Twitter Terhadap Kebijakan Merdeka Belajar”.

2. TINJAUAN PUSTAKA

A. Text Mining

Text mining adalah sebuah penemuan informasi baru namun belum ditemukan, yang secara otomatis diekstraksi dari sumber yang tidak terstruktur (Feldman et al., 2007). Oleh karena itu, *text mining* memerlukan suatu proses yang dapat mengubah data yang tidak terstruktur menjadi data yang lebih terstruktur.

B. Text Preprocessing

Text preprocessing adalah proses memperoleh, memproses, dan mengelola informasi dengan menganalisis bagaimana informasi tersebut berhubungan dengan data yang tidak terstruktur (Suparyanto dan Rosad 2020). Terdapat beberapa langkah, sebagai berikut :

- Cleansing* : Menghapus kata – kata yang tidak perlu seperti karakter HTML, RT (*ReTweet*), *emoticon*, tagar (#), nama pengguna (@), URL (<http://t.o>), email (nama@t.o), simbol dan tanda baca (Rahman Isnain et al., 2021).
- Case Folding* : Menyatukan huruf, menghilangkan karakter non – huruf, dan mengubah semua huruf menjadi huruf kecil (Luqyana, 2019).

- c. *Stopward Removal* : Menghilangkan kata umum yang dianggap tidak mempunyai arti yang penting. Tujuannya adalah mengurangi jumlah kata yang disimpan oleh sistem (Cahyono, 2017).
- d. *Stemming* : Mengubah kata tambahan menjadi kata induk. Ada dua aturan, pertama didasarkan pada pendekatan kamus dan yang kedua pendekatan aturan (Utomo, 2013).
- e. *Normalization* : Proses yang digunakan untuk menyederhanakan proses penguraian kata tidak standar dalam *tweet*. Contohnya, singkatan, tanggal, dan akronim (Aswir & Misbah, 2018).
- f. *Tokenizing* : Memecah kalimat asli menjadi kata dan memberikan pemisah kata seperti titik (.), koma (,), spasi antara kata dan angka, dan karakter lainnya (Rahman Isnain et al., 2021).

C. Analisis Sentimen

Analisis sentimen adalah bagian dari area penelitian seperti pemrosesan bahasa alami dan penambangan data. Tujuan analisis sentimen ini adalah untuk mengetahui bagaimana perasaan seseorang pada suatu teks yang mereka tulis. Menurut Alsaedi dan Khan (2019), analisis sentimen adalah strategi untuk memeriksa penilaian seseorang atau kelompok, seperti halnya dengan investigasi konten, analisis sentimen juga dapat mengetahui pendapat pengguna tentang berbagai topik.

D. Twitter

Twitter adalah suatu layanan *microblogging* resmi diluncurkan pada 13 Juli 2006. Aktivitas umum di twitter yaitu untuk memposting suatu yang biasa disebut "*tweet*". Panjang maksimum "*tweet*" adalah 280 karakter (Fitri et al., 2019). Menurut penelitian dari Bond High Plus (Himalay, 2021) setiap menit pada tahun 2021, 200.000 orang akan men "*tweet*" melalui aplikasi twitter dengan tingkat aktivitas tinggi di Internet.

E. Kebijakan Merdeka Belajar

Kebijakan adalah keputusan yang dibuat oleh pejabat pemerintah dan biasanya digunakan untuk memilih dan menunjukkan kunci keputusan yang mempengaruhi kehidupan

organisasi dan individu pemerintahan. Sebuah kebijakan harus bebas, seperti halnya kebijakan merdeka belajar (Robih, 2019). Menurut Prasetyo (2021) semua lembaga pendidikan menjamin siswa dan guru memiliki kebebasan untuk belajar secara mandiri, mengajar dan berpikir kreatif dalam persaingan di Era Revolusi Industri 4.0 yang semakin meningkat.

F. Term Frequency – Inverse Document Frequency (TF-IDF)

TF-IDF adalah metode yang digunakan untuk menghitung nilai bobot untuk setiap kata dalam sebuah dokumen. Nilai bobot adalah nilai pentingnya kata dalam dokumen. Semakin tinggi bobotnya, semakin penting peran kata dalam dokumen. Secara umum langkah untuk menghitung bobot kata menggunakan *TF-IDF* adalah sebagai berikut (Havrlant & Kreinovich, 2014):

$$W_{dt} = tf_d \times idf_t = tf_d \times \log\left(\frac{N}{df_t}\right)$$

Keterangan :

- W_{dt} = Nilai *term* ke-*t* pada dokumen *d*
- tf_d = Jumlah *term* *t* pada dokumen *d*
- df_t = Jumlah dok mengandung *term* *t*
- N = Jumlah dok secara keseluruhan

G. Multinomial Naïve Bayes (MNB)

Multinomial Naïve Bayes dikembangkan dari algoritma *bayesian* untuk digunakan dalam klasifikasi dokumen atau teks. *Multinomial* memperhitungkan frekuensi dari setiap kata yang muncul dalam dokumen tertentu. Metode ini merupakan salah satu cabang dari algoritma *naïve bayes*, yang memperhitungkan jumlah kemunculan kata dalam dokumen (Kurniawan et al., 2017). Rumus *Multinomial Naïve Bayes*, yaitu (Rahman & Doewes, 2017).

Untuk menentukan peluang awal (probabilitas *prior*) kelas *c*, yaitu :

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

Keterangan :

- N_c = Total kelas *c* pada semua dokumen
- N = Total semua dokumen

Untuk menghitung probabilitas kata ke – n pada kelas c yang digunakan dalam bobot kata *TF-IDF* menggunakan rumus berikut, yaitu :

$$P(t_n|c) = \frac{W_{ct+1}}{(\sum W' \in VW'_{ct+B'})}$$

Keterangan :

- W_{ct} = Nilai *TF-IDF* pada kategori c
- $\sum W' \in VW'_{ct}$ = Total W pada kategori c
- B' = Nilai *idf* yang tidak dikali *tf*

Untuk menghitung probabilitas sebuah dokumen, yaitu :

$$P(c|d) = P(c) \times P(t_1|c) \times P(t_2|c) \times P(t_n|c)$$

Keterangan :

- $P(c|d)$ = Probabilitas dokumen kelas c
- $P(c)$ = Probabilitas *prior* pada kelas c
- $P(t_n|c)$ = Probabilitas kata ke – n kelas c
- t_n = Kata ke – n pada dokumen

H. K-Fold Cross Validation

K-Fold cross validation digunakan untuk menunjukkan tingkat keberhasilan rata – rata dari model *machine learning* yang dibuat. Metode ini menguraikan data menjadi k bagian, masing – masing berisi jumlah data yang seimbang (Zannuar & Lhaksmana, 2021). Data yang dihasilkan berupa data latih dan data uji, akurasi yang diperoleh untuk setiap k iterasi dilakukan untuk memberikan akurasi model yang dibuat.

K-Fold cross validation adalah metode validasi yang membagi k subset dan mengulangi k iterasi untuk pelatihan dan pengujian. Dalam setiap iterasi, satu subset digunakan sebagai data uji dan subset lainnya digunakan sebagai data latih. Data untuk metode *k-fold cross validation* dipartisi menjadi k himpunan bagian dengan ukuran yang sama.

I. Confusion Matrix

Confusion matrix adalah bagaimana perhitungan akurasi dilakukan, *confusion matrix* memberikan keputusan yang dibuat pada data latih dan data uji, yang berisi informasi aktual dan prediksi tentang sistem klasifikasi. Berikut adalah tabel *confusion matrix* :

Tabel 1. Confusion Matrix

Kelas Prediksi	Kelas Aktual	
	Positive (+)	Negative (-)
Positive (+)	True Positive	False Positive
Negative (-)	False Negative	True Negative

Sumber: Jurnal

Keterangan:

- a. *True Positive* adalah banyaknya kelas positif dan diklasifikasikan benar.
- b. *True Negative* adalah banyaknya kelas negatif yang diklasifikasikan benar.
- c. *False Negative* adalah banyaknya kelas negatif yang diklasifikasikan dalam kelas positif.
- d. *False Positive* adalah banyaknya kelas positif yang diklasifikasikan dalam kelas negatif.

J. Uji Performansi

Pengujian performansi ini dijalankan pada nilai yang dihasilkan dari kelas pada prediksi dan nilai asli dari kelas tersebut.

- a. Akurasi, digunakan untuk pengujian yang berdasarkan kedekatan dengan nilai aktual dan nilai prediksi, rumus nilai akurasi dihitung dengan cara (Makhmudah, 2019).

$$\frac{TP + TN}{TP + FP + FN + TN} \times 100\%$$

Keterangan :

- TP = True Positive
- TN = True Negative
- FP = False Positive
- FN = False Negative

- b. Presisi, digunakan untuk menghitung perbandingan dari nilai total informasi yang tercatat oleh sistem baik itu relevan maupun tidak relevan, rumus nilai presisi dihitung dengan cara (Makhmudah, 2019).

$$\frac{TP}{TP + FP} \times 100\%$$

Keterangan :

- TP = True Positive
- FP = False Positive

c. *Recall*, adalah metode penarikan dokumen yang dianggap relevan dan diakui di antara semua dokumen terkait, rumus nilai *recall* dihitung dengan cara (Makhmudah, 2019).

$$\frac{TP}{TP + FN} \times 100\%$$

Keterangan :

TP = True Positive

FN = False Negative

K. Python

Python adalah bahasa pemrograman *open source*, bahasa pemrograman *python* mudah dipahami karena memiliki *sintaks* yang jelas, berisi banyak modul dan *library* yang sudah siap pakai (Harismawan, 2017). Pada bidang analisis data kode yang biasa digunakan seperti, *pandas*, *numpy*, *matplotlib*, *scikit learn*, dan lain sebagainya.

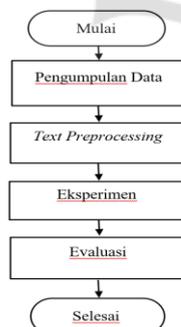
L. Google Collabs

Google collabs adalah produk dari *Google Research*. *Collabs* adalah dokumen yang dapat dieksekusi untuk menulis, menyimpan, dan berbagi program. *Google collabs* adalah *runtime* berbasis *cloud* yang berjalan *browser*, yang memungkinkan pengguna untuk memanfaatkan *library python* untuk analisis dan visualisasi data. Pengguna dapat mengimpor data dari akun *google drive* ke *notebook collabs*.

3. METODOLOGI PENELITIAN

A. Rancangan Penelitian

Desain penelitian memiliki tahapan penting yang akan dilakukan dalam penelitian, agar pelaksanaannya terstruktur dan sistematis dengan baik. Tahapan penelitian analisis sentimen ini adalah :



Gambar 1. Diagram Penelitian

Sumber: Diagram Penelitian

B. Pengumpulan Data

Pengumpulan data pada penelitian ini menggunakan *twitter API (Application Programming Interface)* untuk memperoleh data. Data yang dikumpulkan sebanyak 650 data dari bulan Juni sampai Agustus 2022 dengan menggunakan kata kunci “merdeka belajar”.

C. Pelabelan Data

Data yang telah didapatkan akan melalui proses pelabelan data, dimana data tersebut akan diberikan label berupa variabel positif dan negatif yang dilakukan secara manual dan di validasi oleh salah satu guru Bahasa Indonesia yang berkompeten dalam bidangnya.

Tabel 2. Hasil Pelabelan Data

Data	Positif	Negatif
Hasil <i>Crawling</i>	325	325
Jumlah	650	

Sumber: Hasil Pelabelan Data

4. HASIL DAN PEMBAHASAN

A. Text Preprocessing

Pada tahap pra proses data terdapat beberapa tahapan yaitu, *case folding*, *stopword removal*, *cleansing*, *stemming*, *normalization*, dan *tokenizing*.

B. Pembobotan Kata

Pada proses ini dilakukan untuk mendapatkan nilai kata pada setiap dokumen sentimen, agar dapat dibentuk dalam vektor pada proses pengklasifikasian. Pada proses pembobotan kata, metode yang digunakan adalah *TF-IDF*.

C. Pembagian Data

Proses pembagian data ini dilakukan menggunakan *k-fold cross validation*. Pada penelitian ini nilai *k-fold* yang digunakan terdiri dari *2-fold*, *5-fold*, dan *10-fold*. Dengan membagi data menjadi bagian yang disebut sebagai data uji dan data latih.

D. Hasil Analisis

Setelah melakukan pengujian pada data kebijakan merdeka belajar dengan menggunakan algoritma *Multinomial Naïve Bayes*, maka

berikut adalah tabel hasil dari akurasi, presisi, dan *recall*.

Tabel 3. Hasil Nilai Akurasi, Presisi, dan *Recall*

<i>K-Fold</i>	Langkah Uji	Akurasi	Presisi	<i>Recall</i>
2-Fold	Langkah Uji 1	63.38%	72.15%	60.32%
	Langkah Uji 2	66.46%	62.28%	69.33%
	Rata – Rata	64.92%	64.83%	67.21%
5-Fold	Langkah Uji 1	66.15%	70.15%	66.20%
	Langkah Uji 2	64.62%	75.47%	54.79%
	Langkah Uji 3	70.77%	71.64%	71.64%
	Langkah Uji 4	67.69%	67.21%	65.08%
	Langkah Uji 5	64.62%	70.13%	70.13%
	Rata - Rata	66.77%	65.57%	70.92%
10-Fold	Langkah Uji 1	66.15%	71.43%	67.57%
	Langkah Uji 2	64.62%	65.62%	63.63%
	Langkah Uji 3	61.54%	67.86%	54.29%
	Langkah Uji 4	70.77%	72.00%	60.00%
	Langkah Uji 5	75.38%	76.32%	80.56%
	Langkah Uji 6	64.62%	62.07%	60.00%
	Langkah Uji 7	73.85%	81.48%	64.71%
	Langkah Uji 8	66.15%	64.71%	68.75%
	Langkah Uji 9	56.92%	60.53%	63.89%
	Langkah Uji 10	66.15%	79.49%	68.89%
Rata - Rata	66.62%	65.23%	70.15%	

Sumber: Hasil Perhitungan

Berdasarkan dari keseluruhan nilai yang telah diperoleh seperti tabel 4.12, maka nilai akurasi tertinggi diperoleh pada langkah uji 5 dengan *fold* K=10 sebesar 75.38%, dengan nilai presisi sebesar 76.32%, dan *recall* sebesar 80.56%.

Tabel 4. Hasil *Fold* K=10 Langkah Uji 5

Kriteria	Jumlah
<i>True Positive (TP)</i>	29
<i>True Negative (TN)</i>	20
<i>False Positive (FP)</i>	9
<i>False Negatif (FN)</i>	7

Sumber: Hasil Perhitungan

$$1. \text{ Akurasi} = \frac{TP+TN}{TP+FP+FN+TN} \times 100\%$$

$$= \frac{29+20}{29+9+7+20} \times 100\%$$

$$= 75.38\%$$

$$2. \text{ Presisi} = \frac{TP}{TP+FP} \times 100\%$$

$$= \frac{29}{29+9} \times 100\%$$

$$= 76.32\%$$

$$3. \text{ Recall} = \frac{TP}{TP+FN} \times 100\%$$

$$= \frac{29}{29+7} \times 100\%$$

$$= 80.56\%$$

5. KESIMPULAN DAN SARAN

A. Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan pada “Implementasi Algoritma *Multinomial Naïve Bayes* untuk Analisis Sentimen Twitter Terhadap Kebijakan Merdeka Belajar”, maka kesimpulan yang dapat di ambil adalah :

Hasil akurasi tertinggi diperoleh pada langkah uji 5 dengan *fold* K=10 sebesar 75.38%, dengan nilai presisi sebesar 76.32%, dan *recall* sebesar 80.56%.

B. Saran

Pada penelitian ini tentunya masih terdapat kekurangan, sehingga perlu adanya penyempurnaan oleh penelitian selanjutnya, antara lain :

1. Diharapkan untuk peneliti selanjutnya agar menggunakan jumlah data yang lebih banyak, untuk mendapatkan hasil akurasi yang lebih maksimal.
2. Dapat membandingkan dengan algoritma yang lain untuk mencari hasil yang lebih baik.

6. DAFTAR PUSTAKA

- Alsaeedi, A., & Khan, M. Z. (2019). A study on sentiment analysis techniques of Twitter data. *International Journal of Advanced Computer Science and Applications*, 10(2), 361–374.
<https://doi.org/10.14569/ijacsa.2019.0100248>
- Aswir, & Misbah, H. (2018). No 主観的健康感を中心とした在宅高齢者における健康関連指標に関する共分散構造分析

- Title. *Photosynthetica*, 2(1), 1–13.
<http://link.springer.com/10.1007/978-3-319-76887-8-3-319-93594-2>
<http://dx.doi.org/10.1016/B978-0-12-409517-5.00007-3>
<http://dx.doi.org/10.1016/j.jff.2015.06.018>
<http://dx.doi.org/10.1038/s41559-019-0877-3>
- Cahyono, Y. (2017). Analisis Sentiment pada Sosial Media Twitter Menggunakan Naïve Bayes Classifier dengan Feature Selection Particle Swarm Optimization dan Term Frequency. *Jurnal Informatika Universitas Pamulang*, 2(1), 14. <https://doi.org/10.32493/informatika.v2i1.1500>
- Fitri, V. A., Andreswari, R., & Hasibuan, M. A. (2019). Sentiment analysis of social media Twitter with case of Anti-LGBT campaign in Indonesia using Naïve Bayes, decision tree, and random forest algorithm. *Procedia Computer Science*, 161, 765–772. <https://doi.org/10.1016/j.procs.2019.11.181>
- 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, dan Java. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 2(1), 237–245. <https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/781>
- Havrlant, L., & Kreinovich, V. (2014). A Simple Probabilistic Explanation of Term Frequency-Inverse A Simple Probabilistic Explanation of Term Frequency-Inverse Document Frequency (tf-idf) Heuristic (and Variations Motivated Document Frequency (tf-idf) Heuristic (and Variations Motivated by Th. https://scholarworks.utep.edu/cs_techrep849
https://scholarworks.utep.edu/cs_techrep/849
- Kurniawan, B., Fauzi, M. A., & Widodo, A. W. (2017). Klasifikasi Berita Twitter Menggunakan Metode Improved Naïve Bayes. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer (J-PTIHK) Universitas Brawijaya*, 1(10), 1193–1200.
- Luqyana, T. (2019). Pembangunan Aplikasi Pencarian Produk Dari Video Make Up Tutorial Memanfaatkan Api Youtube Dan Google Speech To Text Berbasis Android. July, 1–23.
- Makhmudah, U. (2019). Analisis Sentimen Terhadap Tweet Kaum Homoseksual Indonesia Menggunakan Metode Support Vector Machine. In *Repository.Unej.Ac.Id*. <https://repository.unej.ac.id/handle/123456789/93827>
<https://repository.unej.ac.id/bitstream/handle/123456789/93827/UmrohMakhmudah-152410101097Sdh.pdf?sequence=1&isAllowed=y>
- 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.
- Rahman, A., & Doewes, A. (2017). Online News Classification Using Multinomial Naive Bayes. *ITSMART: Jurnal Ilmiah Teknologi Dan Informasi*, 6(1), 32–38.
- Rahman Isnain, A., Indra Sakti, A., Alita, D., & Satya Marga, N. (2021). Sentimen Analisis Publik Terhadap Kebijakan Lockdown Pemerintah Jakarta Menggunakan Algoritma Svm. *Jdmsi*, 2(1), 31–37. <https://t.co/NfhmfMjtXw>
- Suparyanto dan Rosad (2015). (2020). 濟無No Title No Title No Title. In *Suparyanto dan Rosad (2015)* (Vol. 5, Issue 3).
- Utomo, M. S. (2013). Implementasi Stemmer Tala pada Aplikasi Berbasis Web. *Jurnal Teknologi Informasi DINAMIK*, 18(1), 41–45. <https://www.unisbank.ac.id/ojs/index.php/fti1/article/view/1673>

Zannuar, M. S., & Lhaksana, K. M. (2021).
*Prediksi Retweet Berdasarkan Feature
User-Based Menggunakan Metode
Klasifikasi Random Forest*. 8(5), 11183–
11191.

