

# Analisis Sentimen Tiktok Shop Menggunakan Metode Multinomial Naïve Bayes Dan BM25

M. Andrew Arjunanda Yasin<sup>1</sup>, Dwi Arman Prasetya<sup>2</sup>, Tresna Maulana Fahrudin<sup>3</sup>

<sup>1</sup>Universitas Pembangunan Nasional Veteran Jawa Timur

<sup>2</sup>Universitas Pembangunan Nasional Veteran Jawa Timur

<sup>3</sup>Universitas Pembangunan Nasional Veteran Jawa Timur

<sup>1</sup>20083010014@student.upnjatim.ac.id, <sup>2</sup>arman.prasetya.sada@upnjatim.ac.id,

<sup>3</sup>tresna.maulana.ds@upnjatim.ac.id

**ABSTRAK.** Dalam era digital saat ini, TikTok merupakan aplikasi populer di kalangan pengguna internet. Salah satu fitur yang menonjol adalah TikTok Shop yang memungkinkan pengguna untuk berbelanja langsung melalui platform tersebut. Namun, pada 4 Oktober 2023 TikTok Shop dihentikan sementara oleh pemerintah karena kebijakan regulasi perdagangan online. Setelah dibuka kembali pada 12 Desember 2023, beragam tanggapan masyarakat muncul sebagai respons terhadap fenomena ini. Penelitian ini melakukan analisis sentimen terhadap TikTok Shop dengan mengamati tanggapan masyarakat di Twitter menggunakan metode klasifikasi Multinomial Naïve Bayes dengan pembobotan kata menggunakan metode BM25 dan TF-IDF. Hasil analisis menunjukkan mayoritas sentimen positif terkait TikTok Shop yang menunjukkan ketidaksetujuan terhadap kebijakan penutupan sebelumnya. Pada pengujian dengan ukuran sampel uji sebesar 10% BM25 menunjukkan kinerja sedikit lebih baik daripada TF-IDF.

**Kata Kunci:** *TikTok Shop; Analisis Sentimen; BM25; Multinomial Naïve Bayes*

**ABSTRACT.** In the current digital era, TikTok has emerged as a popular application among internet users. One prominent feature is the TikTok Shop, which allows users to shop directly through the platform. However, on October 4, 2023, TikTok Shop was temporarily suspended by the government due to online trade regulation policies. After reopening on December 12, 2023, various responses from the public emerged in reaction to this phenomenon. This research conducted sentiment analysis on TikTok Shop by observing public responses on Twitter using the Multinomial Naïve Bayes classification method with word weighting using the BM25 and TF-IDF methods. The analysis results showed a majority of positive sentiment regarding TikTok Shop, indicating disapproval of the previous closure policy. In testing with a 10% test sample size, BM25 showed slightly better performance than TF-IDF.

**Keywords:** *TikTok Shop; Sentiment Analysis; BM25; Multinomial Naïve Bayes*

## PENDAHULUAN

Teknologi informasi telah menjadi bagian tak terpisahkan dalam kehidupan sehari-hari di era digital yang mencakup semua lapisan masyarakat dan industri bisnis. Aplikasi TikTok sangat populer dan merepresentasikan tren tersebut dengan pertumbuhan pengguna yang terus meningkat (Databooks, 2023). Aplikasi ini dilengkapi dengan berbagai filter dan musik sebagai tambahan fitur (Utami, 2021). Saat ini TikTok memiliki fitur inovatif yang dikenal sebagai TikTok Shop yang merupakan inovasi dalam dunia bisnis online. TikTok Shop dapat memudahkan pengguna untuk berbelanja langsung melalui aplikasi TikTok (Yuniarta, 2023).

Namun, operasional TikTok Shop ditutup pada hari Rabu, 4 Oktober 2023. ByteDance selaku pemilik TikTok menutup TikTok Shop untuk mematuhi Peraturan Menteri Perdagangan Nomor 31 Tahun 2023 tentang perizinan Berusaha, Periklanan, Pembinaan, dan Pengawasan Pelaku Usaha dalam Perdagangan melalui Sistem Elektronik. Hingga akhirnya pada Harbolnas 12 Desember 2023 TikTok Shop kembali dibuka. Namun, meskipun dibuka kembali transaksi belanja di TikTok Shop masih dilakukan di dalam satu platform yang sama (DPR, 2023).

Penutupan dan pembukaan kembali TikTok Shop menciptakan respons masif dari masyarakat. Banyak orang memberikan tanggapan dan menyuarakan pendapat mereka melalui media sosial untuk berbagi masukan dan pandangan terkait kejadian ini. Keanekaragaman pesan ini mendorong perlunya menganalisis sentimen atau melakukan penambangan opini. Analisis sentimen membantu memahami apakah pesan-pesan memiliki sifat positif atau negatif dalam jumlah besar yang sulit dilakukan secara manual (Septian, Fachrudin, & Nugroho, 2019).

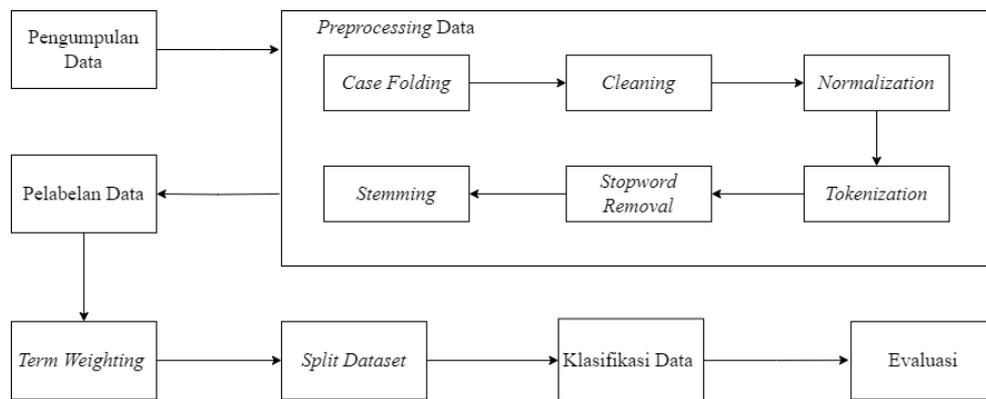
Salah satu metode klasifikasi sentimen yang sering digunakan adalah Naïve Bayes, sebuah algoritma sederhana yang menghitung probabilitas berdasarkan data (Husnina, Ratnawati, & Rahayudi, 2023). Naïve Bayes memiliki beberapa algoritma utama, yaitu Multinomial, Gaussian, dan Bernoulli. Di antara ketiganya, Multinomial Naïve Bayes dianggap sebagai algoritma khusus yang paling efektif dalam klasifikasi (Bahary, Sibaroni, & Mubarak, 2019). Dalam Multinomial Naïve Bayes, setiap kata dalam teks dianggap sebagai fitur

yang independen satu sama lain. Sebelum menerapkan algoritma Multinomial Naïve Bayes untuk mengklasifikasikan data, terdapat langkah penting yang harus dilakukan yaitu pembobotan kata. Salah satu metode yang efektif untuk melakukan ini adalah menggunakan BM25 sebuah metode pemeringkatan yang mengurutkan dokumen berdasarkan seberapa relevan mereka dengan kata kunci dalam pencarian (Frinta, Indriati, & Adikara, 2019). Meskipun ada beberapa metode lain seperti TF-IDF, penelitian (Indriati, Rahayudi, & Dewi, 2021) menunjukkan bahwa pembobotan TF-IDF cenderung memiliki hasil yang lebih rendah dibandingkan dengan BM25.

Dalam penelitian ini, penulis akan membandingkan metode BM25 dan TF-IDF untuk memberikan bobot pada kata-kata dalam teks. Kemudian algoritma Multinomial Naïve Bayes akan digunakan untuk mengklasifikasikan sentimen terhadap TikTok Shop.

## METODE PENELITIAN

Untuk mendukung penulisan maka diperlukan tahapan penelitian yang merupakan langkah-langkah dari proses penelitian. Berikut tahapan penelitian dapat dilihat pada Gambar 1. di bawah ini:



**Gambar 1.** Diagram Alir Penelitian

Berikut ini uraian diagram alir penelitian:

### Pengumpulan Data

Pengumpulan Data pada penelitian ini menggunakan Web Crawling. Web Crawling adalah metode pengumpulan data di mana sebuah program otomatis menjelajahi berbagai halaman web dengan memasukkan URL sebagai acuan (Algifari Rismawan & Syahidin, 2023). Dalam penelitian ini, data diambil dari Tweet menggunakan teknik Web Crawling dengan menggunakan pustaka Python bernama "Tweet-harvest". Pengumpulan data difokuskan pada Bahasa Indonesia untuk memastikan relevansi dan kualitas data terkait peristiwa TikTok Shop di Indonesia.

### Preprocessing Data

Preprocessing Data adalah langkah penting untuk menyiapkan data teks sebelum dilakukan analisis lebih lanjut (Balya, 2019). Proses ini bertujuan menghasilkan data yang lebih baik, terstruktur, dan mudah diolah sehingga analisis dapat dilakukan dengan tingkat akurasi dan efisiensi yang optimal (Al Haromainy, Prasetya, & Sari, 2023).

Proses Preprocessing terdapat beberapa tahapan yang harus diikuti:

1. **Case Folding:** Tahap ini dilakukan untuk mengubah huruf kapital menjadi huruf kecil pada kalimat Tweet. Tujuannya agar teks secara seragam dan konsisten dalam proses analisis.
2. **Cleaning:** Tahap ini dilakukan untuk menghilangkan komponen-komponen yang tidak diperlukan seperti URL, nama akun atau mention, tagar, angka, tanda baca, emotikon, huruf tunggal, kata berulang yang berdekatan, spasi berlebih, baris baru, dan spasi di awal dan akhir teks. Tujuannya agar teks siap untuk analisis lebih lanjut dengan menghilangkan elemen-elemen yang tidak diperlukan.
3. **Normalization:** Tahap ini kata-kata slang dalam teks diganti dengan kata formal yang sesuai berdasarkan daftar normalisasi yang telah dibuat. Tujuannya untuk mengubah kata-kata informal menjadi baku agar teks lebih terstruktur dan konsisten untuk analisis lebih lanjut.
4. **Tokenization:** Tahap ini teks dibagi menjadi token atau kata-kata individual. Tujuannya untuk mempersiapkan teks agar dapat diolah lebih lanjut dengan memisahkan kata-kata menjadi unit terpisah.

5. Stopword Removal: Tahap ini melibatkan penghapusan kata-kata stopwords dari teks yang telah ditokenisasi. Tujuannya untuk membersihkan teks dari kata-kata yang umumnya tidak memberikan informasi penting dalam analisis.
6. Stemming: Tahap ini dilakukan proses Stem pada setiap token kecuali kata-kata yang telah ditentukan. Tujuannya untuk mengubah kata-kata menjadi bentuk dasarnya.

### Pelabelan Data

Proses pelabelan data dalam penelitian ini dilakukan secara manual oleh penulis, setiap Tweet diberi label positif atau negatif berdasarkan sentimen yang terkandung di dalamnya. Selanjutnya, hasil pelabelan tersebut dibahas bersama rekan penulis untuk mengurangi bias dan meningkatkan kualitas data yang digunakan.

### Term Weighting

Term Weighting atau Pembobotan kata merupakan pendekatan untuk menilai pentingnya sebuah kata berdasarkan frekuensinya dalam dokumen. Metode BM25 atau Best Match 25 adalah metode yang digunakan untuk menampilkan dokumen yang relevan dengan menerapkan mekanisme pemeringkatan berdasarkan tingkat kemiripan isi dari query yang diberikan (Pranata, 2019). BM25 dikembangkan oleh Stephen Robertson dan Karen Sparck Jones. Konsep dasarnya melibatkan tiga faktor utama dalam pembobotan kata: Term Frequency (TF), Inverse Document Frequency (IDF), dan nilai rata-rata panjang dokumen (Purnamawati, Winarto, & Mailasari, 2023).

Persamaan BM25 untuk menghitung nilai pemeringkatan ditunjukkan sebagai berikut:

$$\text{Score}(D, Q) = \sum_{i=1}^N \text{IDF}(q_i) \frac{\text{TF}(q_i, d_j) \cdot (k + 1)}{\text{TF}(q_i, d_j) + k \left(1 - b + b \frac{|d_j|}{L}\right)} \quad (1)$$

Keterangan:

$\text{Score}(D, Q)$  = Score kemiripan query dengan dokumen

$\text{IDF}(q_i)$  = Inverse Document Frequency  $q_i$

$\text{TF}(q_i, d_j)$  = Term Frequency  $q_i$  pada dokumen  $d_j$

$L$  = Nilai rata-rata jumlah dokumen

$N$  = Jumlah dokumen

$|d_j|$  = banyaknya kata dalam dokumen  $d_j$

$k = 1,2 \leq k \leq 2,0$

$b = 0,5 \leq b \leq 0,8$

### Split Dataset

Dalam penelitian ini, pembagian data menggunakan metode holdout dengan membagi Dataset menjadi dua bagian utama: data latih (training) dan data uji (testing). Data latih berperan sebagai tempat untuk melatih model klasifikasi agar model dapat mempelajari pola dan karakteristik dari data yang sudah memiliki label sentimen. Setelah melalui tahap pelatihan tersebut model dapat diuji menggunakan data uji untuk mengukur seberapa baik performanya dalam mengklasifikasikan data baru yang belum pernah dilihat sebelumnya.

### Klasifikasi Data

Klasifikasi data adalah proses pengelompokan data berdasarkan atribut atau kategori sentimennya menggunakan model klasifikasi yang didasarkan pada pengetahuan yang diperoleh dari data pelatihan. Dalam penelitian ini, klasifikasi dilakukan menggunakan algoritma Multinomial Naïve Bayes. Multinomial Naïve Bayes adalah variasi dari metode Naïve Bayes yang menghitung frekuensi setiap kata pada dokumen (Oktavia, 2023). Metode ini dirancang untuk menentukan frekuensi kata atau Term dalam suatu dokumen. Karena kata-kata dapat menjadi kunci dalam menentukan sentimen sehingga membuat Multinomial Naïve Bayes sering

digunakan dalam analisis sentimen (Singh, Kumar, Gaur, & Tyagi, 2019). Probabilitas sebuah dokumen d berada di kelas c dapat dihitung dengan rumus:

$$P(c|d) \propto P(c) \prod_{1 \leq k \leq nd} P(t_k|c) \tag{2}$$

Keterangan:

$P(t_k|c)$  adalah probabilitas kondisional dari kata  $t_k$  dalam dokumen dari kelas c.

$P(c)$  adalah probabilitas prior dari dokumen dalam kelas c.

**Evaluasi**

Untuk menilai seberapa baik kinerja model klasifikasi yang telah dibangun diperlukan beberapa metrik evaluasi. Confusion Matrix memberikan metrik seperti akurasi, presisi, recall, dan f-measure yang dapat membantu memahami seberapa baik model dalam mengklasifikasikan data (Normawati & Prayogi, 2021).

**HASIL DAN PEMBAHASAN**

Pada bagian ini penulis akan membahas hasil dari metode penelitian.

**Pengumpulan Data**

Penulis mengumpulkan data dari tanggal 6 September 2023 hingga 22 Januari 2024 dari *Tweet* dengan beberapa keyword yaitu Penutupan TikTok Shop dan Pembukaan TikTok Shop serta dari komentar yang membahas tentang dua kejadian diatas. Terdapat beberapa kolom yaitu *created\_at*, *id\_str*, *full\_text*, *quote\_count*, *reply\_count*, *reTweet\_count*, *favorite\_count*, *lang*, *user\_id\_str*, *conversation\_id\_str*, *username*, *Tweet\_url*. Namun, penulis hanya akan fokus pada kolom *full\_text* dan data yang didapatkan hingga 1086 baris.

**Preprocessing Data**

Preprocessing Data digunakan untuk memastikan bahwa data yang akan diolah dalam penelitian telah dibersihkan dan siap untuk diproses. Berikut ini adalah hasil tahapan-tahapannya:

**Tabel 1.** Hasil Preprocessing Data

Full_Text	Case Folding	Cleaning	Normalization	Tokenization	Stopword Removal	Stemming
Mau liburan tapi selama kagak ada pendapatan pribadi rasanya kok sayang uang ya wkwk gara2 tiktok shop tutup :(	Mau. liburan tapi selama kagak ada pendapatan pribadi rasanya kok sayang uang ya wkwk gara2 tiktok shop tutup :(	mau liburan tapi selama kagak ada pendapata n pribadi rasanya kok sayang uang ya wkwk gara tiktok shop tutup	ingin liburan namun selama tidak ada penghasilan pribadi rasanya kok sayang uang wkwk karena tiktok shop tutup	['ingin', 'liburan', 'namun', 'selama', 'tidak', 'ada', 'penghasilan', 'pribadi', 'rasanya', 'kok', 'sayang', 'uang', 'wkwk', 'karena', 'tiktok', 'shop', 'tutup']	['liburan', 'penghasila n', 'pribadi', 'sayang', 'uang', 'wkwk', 'tiktok', 'shop', 'tutup']	['libur', 'hasil', 'pribadi', 'sayang', 'uang', 'wkwk', 'tiktok', 'shop', 'tutup']

**Pelabelan Data**

Pelabelan data dilakukan dalam dua sentimen, yaitu positif dan negatif. Tweet diberi label positif jika mereka menyukai fitur TikTok Shop, menentang penutupan TikTok Shop, dan menyambut pembukaan kembali TikTok Shop. Sebaliknya, Tweet diberi label negatif jika mereka tidak menyukai Fitur TikTok Shop menyetujui penutupan TikTok Shop, dan menentang pembukaan kembali TikTok Shop. Dengan demikian, pelabelan bertujuan untuk mengklasifikasikan Tweet berdasarkan tanggapan positif atau negatif terhadap TikTok Shop. Berikut adalah hasil dari pelabelan data:

**Tabel 2.** Hasil Pelabelan Data

Full_Text	Label
Mau liburan tapi selama kagak ada pendapatan pribadi rasanya kok sayang uang ya wkwk gara2 tiktok shop tutup :(	Positif
Tutup permanen aja TikTok Shop	Negatif

Hasil Pelabelan Data didapatkan bahwa label positif lebih banyak daripada positif yaitu 615 sedangkan negatif sebanyak 471 dari total dari 1086 data.

**Term Weighting**

Proses Term Weighting dalam penelitian ini menggunakan metode BM 25 dan TF-IDF. Term Weighting bertujuan untuk menentukan seberapa penting setiap kata dalam dokumen. Berikut ini adalah contoh hasil pembobotan kata dengan BM25 dan TF-IDF:

**Tabel 1.** Hasil Pembobotan Kata BM25 dan TF-IDF

BM25		TF-IDF	
shop	tiktok	shop	tiktok
0.067	0.067	0.064	0.063
0.029	0.029	0.033	0.033
0.034	0.03	0.043	0.043
0.092	0.091	0.078	0.077
0.083	0.083	0.073	0.073
...	...	...	...

Tabel 3. menunjukkan hasil pembobotan kata menggunakan metode BM25 dan TF-IDF. Meskipun kata-kata yang sama dievaluasi, namun nilai bobotnya berbeda antara metode BM25 dan TF-IDF. Lalu, berikut ini adalah kata-kata dengan bobot tertinggi dan terendah untuk setiap pembobotan:

**Tabel 2.** Kata dengan Bobot Tertinggi dan Terendah

BM25		TF-IDF	
Kata	Total Bobot	Kata	Total Bobot
tiktok	121.866214	tiktok	95.660.277
shop	107.979725	shop	88.415.108
tutup	94.421380	tutup	79.525.883
buka	63.888248	buka	53.848.550
jual	42.043575	umkm	37.802.289
...	...	...	...
pozas	0.105508	cascavel	0.173732
mini	0.105508	saff	0.173732
las	0.105508	juara	0.173732
juara	0.105508	solar	0.173732
cascavel	0.105508	pozas	0.173732

Tabel 4. menunjukkan total bobot untuk kata-kata tertentu menggunakan metode BM25 dan TF-IDF. Kata-kata yang sering muncul seperti "tiktok" dan "shop" memiliki total bobot yang lebih tinggi dalam metode BM25 dibandingkan dengan metode TF-IDF, sedangkan kata-kata yang jarang muncul seperti "pozas" dan "juara" memiliki total bobot yang lebih tinggi dalam metode TF-IDF. Perbedaan ini terjadi karena kedua metode tersebut menggunakan pendekatan yang berbeda dalam mengukur signifikansi suatu kata dalam sebuah dokumen. BM25 lebih memperhatikan frekuensi kata serta panjang dokumen secara keseluruhan, sementara

TF-IDF fokus pada keunikkan dan kemunculan kata dalam sebuah dokumen dalam kumpulan data yang lebih besar.

**Klasifikasi Data**

Peneliti melakukan split dataset yang akan digunakan dalam tahap klasifikasi data, dengan perbandingan data train dan data test yaitu, 90:10, 80:20, dan 70:30. Pengujian ini menggunakan metode pembobotan BM25 dan TF-IDF. Di bawah ini adalah hasil klasifikasinya:

**Tabel 3.** Hasil Akurasi

Test_Size	Pembobotan	Akurasi
10	BM25	0.77
	TF-IDF	0.74
20	<b>BM25</b>	<b>0.80</b>
	<b>TF-IDF</b>	<b>0.80</b>
30	BM25	0.76
	TF-IDF	0.76

Pada Tabel 5. BM25 lebih mengungguli TF-IDF. Pada BM25 didapatkan akurasi tertinggi pada Test Size 20 dengan akurasi 0.80 begitu pula dengan TF-IDF. Tetapi, Pada Test Size 10 BM25 mengungguli dengan skor 0.77 sedangkan TF-IDF mendapatkan skor 0.74.

**Evaluasi**

Untuk mengetahui algoritma yang memiliki kinerja terbaik dalam pengklasifikasian sentimen TikTok Shop maka perlu dilakukan perbandingan terhadap evaluasi model klasifikasi. Untuk menilai evaluasi model diperlukan Confusion Matrix. Berikut ini hasil Confusion Matrix:

**Tabel 4.** Hasil Confusion Matrix

TP		TN		FP		FN	
BM25	TF-IDF	BM25	TF-IDF	BM25	TF-IDF	BM25	TF-IDF
118	119	56	55	27	28	17	16

Dari hasil tabel 6. didapatkan nilai akurasi, presisi, recall, f1-score, dll. Berikut ini hasil evaluasi algoritma dengan pembobotan BM25 dan TF-IDF:

**Tabel 5.** Hasil Evaluasi dengan Pembobotan BM25

	precision	recall	f1-score	support
0	0.77	0.67	0.72	83
1	0.81	0.87	0.84	135
accuracy			0.80	218
macro avg	0.79	0.77	0.78	218
weighted avg	0.80	0.80	0.80	218

**Tabel 6.** Hasil Evaluasi dengan Pembobotan TF-IDF

	precision	recall	f1-score	support
0	0.77	0.67	0.72	83
1	0.81	0.87	0.84	135
accuracy			0.80	218
macro avg	0.79	0.77	0.78	218
weighted avg	0.80	0.80	0.80	218

Terdapat sedikit perbedaan dalam Tabel 8. dan 9. terutama pada nilai recall untuk BM25 dan TF-IDF. Untuk label 0 nilai recall pada BM25 adalah 0.67, sedangkan pada TF-IDF adalah 0.66. Sementara untuk label 1 nilai recall pada BM25 adalah 0.87, sedangkan pada TF-IDF adalah 0.88. Perbedaan yang terjadi pada nilai recall antara metode BM25 dan TF-IDF hanya selisih 0.01. Perbedaan terakhir terdapat pada f1-score, BM25 lebih unggul satu angka daripada TF-IDF yaitu 0.72 dan 0.71. Hal ini menunjukkan bahwa keduanya memiliki performa yang hampir sama dalam mengenali tweet yang seharusnya diklasifikasikan ke dalam label 0 (negatif) dan label 1 (positif).

## KESIMPULAN DAN SARAN

Analisis sentimen terhadap TikTok Shop di Twitter menggunakan Metode Multinomial Naïve Bayes dengan BM25 menunjukkan bahwa mayoritas pengalaman yang diungkapkan adalah positif. Hal ini menandakan bahwa banyak dari masyarakat tidak menyetujui penutupan TikTok Shop dan banyak yang menyambut kembalinya fitur TikTok Shop atau bahwa banyak dari mereka menyukai fitur tersebut. Oleh karena itu, penting bagi pemerintah untuk membuat keputusan yang bijaksana agar tidak merugikan banyak pihak. Metode Multinomial Naive Bayes mampu memprediksi sentimen dengan baik, hingga mencapai akurasi sebesar 80%. Metode ini memberikan kinerja terbaik dengan akurasi yang sama baiknya menggunakan kedua jenis pembobotan, baik BM25 maupun TF-IDF. Namun, pada pengujian dengan ukuran 10%, BM25 sedikit lebih unggul daripada TF-IDF. Hal ini menunjukkan bahwa BM25 tidak secara keseluruhan lebih baik dari TF-IDF dalam pembobotan. Oleh karena itu, disarankan untuk mencoba membandingkan BM25 dengan teknik pembobotan lainnya dan menggunakan seleksi fitur untuk meningkatkan akurasinya.

## DAFTAR RUJUKAN

- Al Haromainy, M. M., Prasetya, D. A., & Sari, A. P. (2023). Improving Performance of RNN-Based Models With Genetic Algorithm Optimization For Time Series Data. *TIERS Information Technology Journal*, 4(1), 16–24. <https://doi.org/10.38043/tiers.v4i1.4326>
- Algifari Rismawan, S., & Syahidin, Y. (2023). Implementasi Website Berita Online Menggunakan Metode Crawling Data Dengan Bahasa Pemrograman Python. *Jurnal Teknik Informatika Dan Sistem Informasi*, 10(3), 167–178. Retrieved from <https://doi.org/10.35957/jatisi.v10i3.4902>
- Bahary, A. F., Sibaroni, Y., & Mubarak, M. S. (2019). Sentiment analysis of student responses related to information system services using Multinomial Naïve Bayes (Case study: Telkom University). *Journal of Physics: Conference Series*, 1192(1). <https://doi.org/10.1088/1742-6596/1192/1/012046>
- Balya. (2019). Analisis Sentimen Pengguna Youtube Di Indonesia Pada Review Smartphone Menggunakan Naïve Bayes. *Skripsi Universitas Sumatera Utara*, pp. 4–16.
- Databoks. (2023). Indonesia Punya Pengguna TikTok Terbanyak ke-2 di Dunia. Retrieved March 15, 2024, from <https://databoks.katadata.co.id/datapublish/2023/11/22/indonesia-punya-pengguna-tiktok-terbanyak-ke-2-di-dunia>
- DPR. (2023). Parlemen Terkini - Dewan Perwakilan Rakyat. Retrieved March 15, 2024, from <https://www.dpr.go.id/berita/detail/id/46750/t/Media+Sosial+Kini+Dilarang+Jualan%2C+Legislator+Ingatkan+Pemerintah+Ciptakan+Regulasi+yang+Adil>
- Frinta, K., Indriati, & Adikara, P. P. (2019). Pencarian Berita Berbahasa Indonesia Menggunakan Metode BM25. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 3(3), 2589–2595. Retrieved from <http://j-ptiik.ub.ac.id>
- Husnina, D. N. N., Ratnawati, D. E., & Rahayudi, B. (2023). Analisis Sentimen Pengguna Aplikasi RedBus berdasarkan Ulasan di Google Play Store menggunakan Metode Naïve Bayes. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 7(2), 737–743. Retrieved from <https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/12297%0Ahttps://j-ptiik.ub.ac.id/index.php/j-ptiik/article/download/12297/5593>
- Indriati, I., Rahayudi, B., & Dewi, C. (2021). Analisis Sentimen Mengenai Moda Raya Terpadu (MRT) Jakarta dengan Metode BM25 dan K-Nearest Neighbor. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 8(2), 389–394. <https://doi.org/10.25126/jtiik.2021824508>
- Normawati, D., & Prayogi, S. A. (2021). Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter. *Jurnal Sains Komputer & Informatika (J-SAKTI)*, 5(2), 697–711.
- Oktavia, K. (2023). Pendekatan multinomial Naive Bayes untuk klasifikasi jenis-jenis cyber harassment di media sosial Twitter. Retrieved from <http://etheses.uin-malang.ac.id/40836/%0Ahttp://etheses.uin-malang.ac.id/40836/1/19610079.pdf>
- Pranata, A. I. . (2019). Klasifikasi Dokumen pada Laporan Kepolisian dengan Menggunakan Metode BM25

- dan Improved K-Nearest Neighbor (IKNN).pdf.*
- Purnamawati, A., Winarto, M. N., & Mailasari, M. (2023). Analisis Sentimen Aplikasi TikTok menggunakan Metode BM25 dan Improved K-NN Fitur Chi-Square. *Jurnal Komtika (Komputasi Dan Informatika)*, 7(1), 97–105. <https://doi.org/10.31603/komtika.v7i1.8938>
- Septian, J. A., Fachrudin, T. M., & Nugroho, A. (2019). Analisis Sentimen Pengguna Twitter Terhadap Polemik Persepakbolaan Indonesia Menggunakan Pembobotan TF-IDF dan K-Nearest Neighbor. *Journal of Intelligent System and Computation*, 1(1), 43–49. <https://doi.org/10.52985/insyst.v1i1.36>
- Singh, G., Kumar, B., Gaur, L., & Tyagi, A. (2019). Comparison between Multinomial and Bernoulli Naïve Bayes for Text Classification. *2019 International Conference on Automation, Computational and Technology Management, ICACTM 2019*, (May 2020), 593–596. <https://doi.org/10.1109/ICACTM.2019.8776800>
- Utami, A. D. V. (2021). Aplikasi Tiktok Menjadi Media Hiburan Bagi Masyarakat Dan Memunculkan Dampak Ditengah Pandemi Covid-19. *MEDIALOG: Jurnal Ilmu Komunikasi*, 4(1), 40–47. <https://doi.org/10.35326/medialog.v4i1.962>
- Yuniarta, A. (2023). Minat Konsumen Membeli Produk Online Melalui Tik Tok Shop Ditinjau Dari. *Journal of Student Research*, 1(3), 39–47.