

PENERAPAN ALGORITMA NAIVE BAYES CLASSIFIER UNTUK ANALISIS SENTIMEN PADA ULASAN APLIKASI MAXIM DI GOOGLE PLAY STORE

Satria Brian Yudanto, Dwi Hartanti, Afu Ichsan Pradana

Teknik Informatika, Universitas Duta Bangsa Surakarta
Jl. Bhayangkara No. 55, Tipes, Serengan, Surakarta, Indonesia
220103167@mhs.udb.ac.id

ABSTRAK

Berkembangnya layanan transportasi berbasis online seperti Maxim telah memberikan kemudahan yang signifikan bagi masyarakat, namun berdampak pada menumpuknya ulasan publik di Google Play Store menjadi tidak terstruktur sehingga menyulitkan penyedia layanan dalam mengevaluasi tingkat kepuasan serta persepsi publik terhadap layanan aplikasi tersebut. Oleh sebab itu, penelitian ini bertujuan guna mengklasifikasikan sentimen ulasan pengguna secara efisien dan otomatis. Penyelesaian masalah dilakukan dengan menerapkan metode *Naive Bayes Classifier*. Pengumpulan data melibatkan 2.500 ulasan melalui *web scraping*, dilanjutkan dengan pelabelan semi-otomatis, validasi manual, dan tahap *preprocessing data* termasuk normalisasi bahasa *slang*. Kemudian melakukan pembobotan dengan TF-IDF dan dibagi menjadi data latih dan uji dengan rasio 80:20. Hasil evaluasi menunjukkan bahwa model mendapat hasil yang baik dengan tingkat akurasi sebesar 94,01%, untuk kelas sentimen positif mendapat nilai *precision* 0.94, *recall* 0.99, dan *f1-score* 0.97, sedangkan pada kelas negatif masing-masing sebesar 0.92, 0.65, dan 0.76. Pada visualisasi data dengan diagram batang dan wordcloud menunjukkan bahwa kepuasan mayoritas pengguna disebabkan oleh performa mitra pengemudi, sedangkan keluhan umumnya berkaitan dengan kendala teknis pada aplikasi.

Kata kunci : maxim, naive bayes, sentimen analisis, ulasan

1. PENDAHULUAN

Kemajuan teknologi yang cepat telah memudahkan pengguna dapat mengakses layanan transportasi secara daring (*online*), yang membuat pemesanan dapat dilakukan dengan sangat praktis melalui aplikasi di ponsel pintar [1]. Perkembangan teknologi tersebut membawa dampak yang sangat besar, salah satunya adalah peningkatan popularitas transportasi daring yang kini telah bergeser menjadi sebuah kebutuhan pokok dalam mobilitas masyarakat. Di Indonesia, Maxim hadir sebagai salah satu perusahaan yang menawarkan layanan transportasi daring serta berbagai layanan lainnya.

Aplikasi Maxim dapat diunduh secara gratis oleh pengguna ponsel pintar melalui platform Google Play Store. Namun, tingginya jumlah unduhan dan pengguna aktif memunculkan tantangan baru, yakni menumpuknya opini publik. Pengguna dapat memanfaatkan kolom ulasan di Google Play Store untuk menceritakan pengalaman mereka, yang bentuknya sangat beragam mencakup opini positif maupun sentimen negatif berupa keluhan, kritik, dan saran [2]. Dengan jumlah ulasan yang sangat banyak, kumpulan teks tersebut menjadi tidak terstruktur. Hal ini menimbulkan masalah efisiensi, sebab pihak penyedia layanan akan sangat kesulitan jika harus membaca dan mengevaluasi ulasan tersebut satu per satu secara manual demi mengukur tingkat kepuasan pengguna secara keseluruhan.

Untuk mengatasi masalah tersebut, diperlukan komputasi otomatis yang mampu mengklasifikasikan opini pengguna dengan efisien. Sebab itu, penelitian ini memanfaatkan metode *Naive Bayes Classifier* guna

menganalisis sentimen ulasan aplikasi Maxim dengan tujuan mengklasifikasikan ulasan menjadi kelas sentimen positif dan negatif.

2. TINJAUAN PUSTAKA

2.1. Analisis Sentimen

Analisis sentimen merupakan metode yang berfungsi guna mendapatkan informasi dari teks secara otomatis. Teknik ini biasanya dipakai untuk mengetahui kecenderungan opini publik terhadap suatu topik guna menentukan apakah pandangan tersebut mengarah ke arah positif atau negatif [3].

2.2. Web Scraping

Web scraping merupakan teknik mengambil data dari suatu halaman *web* secara otomatis dengan bantuan perangkat lunak atau *script* tertentu. Tujuan utama dari teknik ini adalah untuk mengambil data dari berbagai situs dan menyusunnya menjadi format yang terstruktur. Data yang telah terkumpul tersebut dapat dimanfaatkan sebagai bahan baku utama untuk keperluan analisis, pemantauan tren, maupun riset [4].

2.3. Text Mining

Text Mining merupakan rangkaian tahap pengolahan teks guna menemukan informasi penting seperti pola, konsep maupun kata kunci yang terkandung dalam data [5]. Penerapan metode ini menjadi solusi yang sangat efektif untuk mengatasi kendala keterbatasan waktu dalam mengolah data ulasan yang sangat besar dan bertambah setiap harinya.

2.4. Preprocessing Data

Preprocessing data merupakan proses pengolahan teks dengan tujuan membersihkan data dari unsur yang tidak dibutuhkan agar dapat digunakan pada klasifikasi [6]. Tahap ini biasanya mencakup *cleaning, case folding, tokenizing, stopword removal, dan stemming*. [7].

2.5. TF-IDF

TF-IDF (*Term Frequency-Inverse Document Frequency*) merupakan teknik pembobotan pada teks yang berfungsi guna mengukur seberapa penting sebuah kata berdasarkan kemunculannya dalam dokumen. Pada metode TF-IDF, kata yang sering muncul dalam suatu dokumen akan memperoleh bobot yang lebih tinggi. Namun, jika kata tersebut juga banyak ditemukan pada dokumen lain, maka bobotnya menjadi lebih rendah [8].

2.6. Naive Bayes

Naive Bayes Classifier merupakan metode klasifikasi yang menggunakan pendekatan peluang. Metode tersebut memiliki dasar yang beranggapan setiap kemunculan kata atau variabel itu saling bersifat independen atau tidak terikat satu sama lain [9] [10]. Rumus untuk probabilitas bersyarat dalam Algoritma *Naive Bayes Classifier* adalah berikut ini:

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)} \quad (1)$$

$P(A|B)$: Probabilitas bersyarat A yang diberikan oleh B

$P(B|A)$: Probabilitas bersyarat B yang diberikan oleh A

$P(A)$: Probabilitas kejadian A

$P(B)$: Probabilitas kejadian B

2.7. Evaluasi Model Klasifikasi

Setelah model klasifikasi diterapkan, maka dilakukan evaluasi agar bertujuan memeriksa apakah model sudah sesuai pada saat mengklasifikasikan data ulasan pengguna. Metode yang digunakan adalah *confusion matrix*, yang berfungsi sebagai alat ukur untuk membandingkan nilai prediksi sistem terhadap nilai aslinya [11]. Performa model dievaluasi melalui empat nilai utama. Pertama, akurasi untuk melihat persentase kebenaran prediksi secara umum. Kedua dan ketiga adalah *precision* dan *recall* yang menilai ketepatan prediksi positif, dan mengukur kemampuan sistem dalam menjaring data yang relevan. Kemudian, *f1-score* untuk menyeimbangkan keduanya agar bisa melihat hasil kerja model secara keseluruhan.

2.8. Confusion Matrix

Confusion Matrix digunakan sebagai alat evaluasi guna mengukur kinerja dan akurasi algoritma klasifikasi pada data pengujian. Metode ini bertujuan agar mengetahui seberapa baik kinerja suatu model

dapat melakukan proses klasifikasi. *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN) adalah hasil prediksi dari *confusion matrix* [12].

2.9. Penelitian Terdahulu

Menurut [2] dalam penelitian yang berjudul “Sentimen Analisis Terhadap Review Aplikasi Maxim di Google Play Store Menggunakan Support Vector Machine” bertujuan menganalisis opini pengguna dengan menggunakan 1.200 data ulasan hasil *scraping* dari aplikasi Maxim dari Google Play Store dan menunjukkan akurasi 79% dengan metode *Support Vector Machine* (SVM) [2].

Menurut [1] dengan penelitian yang berjudul “Analisis Sentimen Terhadap Aplikasi Maxim Menggunakan Algoritma Forest” mengklasifikasikan ulasan aplikasi Maxim di Google Play Store dan Twitter. Metode yang diterapkan adalah menggunakan *Random Forest* dengan validasi *K-Fold Cross Validation* ($k = 10$), mendapatkan tingkat akurasi sebesar 64% [1].

Menurut [13] dalam penelitian yang berjudul “Implementasi *Bidirectional LSTM* untuk Analisis Sentimen Terhadap Layanan Grab Indonesia” yang melakukan pengujian terhadap analisis sentimen pelayanan Grab Indonesia mendapatkan akurasi 91% dengan algoritma *Bidirectional LSTM* dibandingkan menggunakan *LSTM* biasa yang menghasilkan akurasi sebesar 76% [13].

Menurut [14] dalam penelitian yang berjudul “Analisis Algoritma K-Nearest Neighbor terhadap Sentimen Pengguna Aplikasi Shopee” mengklasifikasikan ulasan pengguna aplikasi Shopee dengan *K-Nearest Neighbor* mendapat akurasi, *precision*, serta *recall* sebesar 70%, 50,5%, dan 44,8% [14].

3. METODOLOGI PENELITIAN

3.1. Jenis dan Sumber Data

Data yang digunakan terdiri dari data primer dan data sekunder, data primer berupa kumpulan teks ulasan pengguna aplikasi Maxim yang diambil dari halaman Google Play Store menggunakan teknik *web scraping*. Untuk data sekunder yang digunakan berasal dari berbagai penelitian sebelumnya meliputi buku, ataupun dari jurnal ilmiah dan literatur yang relevan. Data sekunder tersebut digunakan sebagai dasar teori dan acuan terkait analisis sentimen dan algoritma *Naive Bayes Classifier*.

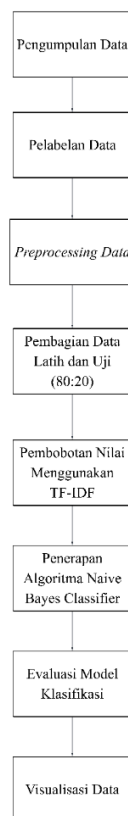
3.2. Tahapan Penelitian

Proses penelitian dimulai dengan mengumpulkan ulasan aplikasi Maxim dari halaman Google Play Store dengan *web scraping*. Setelah ulasan terkumpul, selanjutnya adalah pelabelan data secara semi-otomatis berdasarkan rating. Tahap tersebut dilanjutkan dengan pengecekan kembali secara manual untuk mengoreksi kesalahan data, yaitu memperbaiki label ulasan yang ratingnya tidak relevan

dengan makna teks ulasannya. Untuk rating bintang 3 tidak akan digunakan karena dapat berpotensi menurunkan performa model klasifikasi.

Data yang sudah melewati pelabelan semi-otomatis tadi kemudian masuk ke tahap *preprocessing data* untuk membersihkan teks sebelum diproses lebih lanjut. Tahap ini meliputi menghapus tanda baca, emoji, dan angka (*cleaning*), merubah isi ulasan menjadi huruf kecil (*case folding*). Kemudian melakukan normalisasi *slang* untuk memperbaiki kata singkatan atau kesalahan pengetikan berdasarkan kamus yang telah dibuat [15]. Proses *preprocessing data* dilanjut dengan menggunakan library Sastrawi untuk menghapus kata hubung (*stopword removal*), dan merubah kata berimbuhan menjadi kata dasar (*stemming*), serta terakhir memecah kalimat menjadi kata tunggal (*tokenizing*).

Setelah data dibersihkan, Data kemudian dibagi dengan perbandingan 80:20 untuk data latih serta data uji. Kemudian, dilakukan pembobotan nilai menggunakan TF-IDF guna memberikan nilai pada tiap kata berdasarkan frekuensi kemunculannya. Lalu, *Naive Bayes Classifier* diterapkan guna menentukan kategori sentimen ulasan tersebut. Performa model klasifikasi dievaluasi dengan *confusion matrix* guna mengukur tingkat akurasi, *precision*, *recall*, serta *f1-score*. Terakhir, analisis sentimen yang telah dilakukan divisualisasikan dengan menggunakan diagram batang guna mengetahui top 20 kata yang muncul, serta *wordcloud* guna memvisualisasikan kata dominan yang muncul pada kelas sentimen.



Gambar 1. Tahapan penelitian

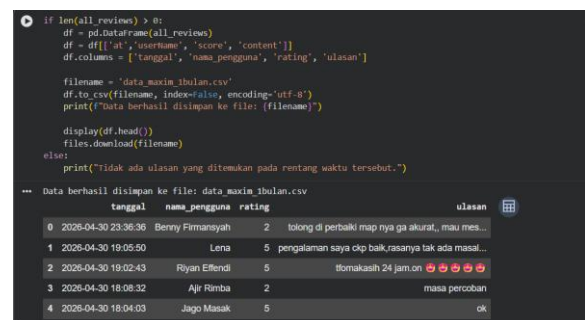
4. HASIL DAN PEMBAHASAN

4.1. Proses Pengumpulan dan Pelabelan Data

Data yang diambil merupakan ulasan aplikasi Maxim dari halaman Google Play Store menggunakan *web scraping*. Ulasan diambil dari rentang 1 bulan kebelakang dengan maksimal 2500 ulasan.

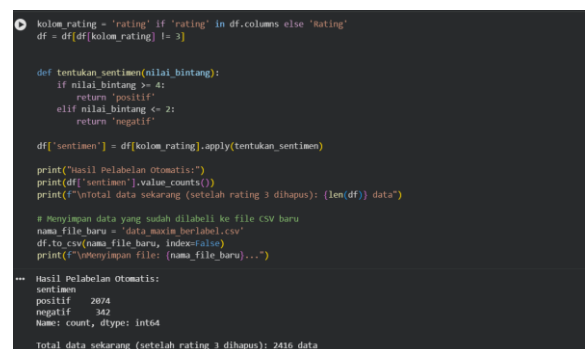
Tabel 1. Contoh hasil ulasan

nama_pengguna	rating	ulasan
Lena	5	pengalaman saya ckp baik,rasanya tak ada masalah,info akun saya SM spt yg tadi
Riyan Effendi	5	tfomakasih 24 jam.on 😊😊😊😊😊
Jago Masak	5	ok
ZAKARIA AL ANSHORI	5	jelas lebih hemat dengan maxim
Umi Kulsum	5	pokok sipp



Gambar 2. Ulasan aplikasi maxim yang diambil menggunakan teknik *web scraping*

Setelah mengumpulkan data, maka data tersebut akan melalui proses pelabelan semi-otomatis dimana rating 4–5 masuk pada ulasan sentimen positif, sementara rating 1–2 masuk pada sentimen negatif. Sementara untuk ulasan dengan rating 3 tidak akan digunakan karena dapat berpotensi menurunkan performa model klasifikasi. Kemudian dilanjutkan dengan pengecekan kembali secara manual untuk memperbaiki ulasan yang tidak relevan dengan ratingnya.



Gambar 3. Proses pelabelan data secara semi-otomatis

user_id	review_text	rating	sentiment
30/04/2020 01:40	Handphone	5	negatif
29/04/2020 14:25	BUDI TRI UTOMO	5	negatif
29/04/2020 13:38	Sybilana Arifiani Putri	5	negatif
29/04/2020 01:47	Very Buram	5	negatif
29/04/2020 00:33	Bara Satun	5	negatif
29/04/2020 01:18	Unika ides	5	negatif
27/04/2020 09:26	Sumana Sumari	5	negatif
27/04/2020 09:23	Fatri Anas	5	negatif
26/04/2020 16:36	Saf Kurnia	5	negatif
26/04/2020 16:30	Nuku Hendra	5	negatif
26/04/2020 09:34	Toko Bangsa Arya Florist Malang	5	negatif
25/04/2020 21:17	Nery Kristiana	5	negatif
24/04/2020 16:15	OKA production	5	negatif
23/04/2020 10:17	Mayanti Yanti	5	negatif
23/04/2020 04:51	Melvi ayu	5	negatif
23/04/2020 15:34	vidvade	5	negatif
22/04/2020 01:37	Falanyus Andia	5	negatif
21/04/2020 16:16	ABD, Halim	5	negatif
21/04/2020 09:08	dian net	5	negatif
20/04/2020 01:14	Karnomata Zahra	5	negatif
19/04/2020 16:50	Israk Raska	5	negatif
18/04/2020 10:13	Romy Augustin	5	negatif
18/04/2020 02:12	Desi Yanti	5	negatif

Gambar 4. Data yang sudah dicek kembali secara manual

4.2. Preprocessing Data

Setelah dilakukan pelabelan secara semi-otomatis, selanjutnya data ulasan melewati proses *preprocessing data* yang meliputi penghapusan tanda baca, emoji, dan angka (*cleaning*), merubah isi ulasan menjadi huruf kecil (*case folding*), normalisasi *slang* untuk memperbaiki kata singkatan atau kesalahan pengetikan berdasarkan kamus yang telah dibuat, dan menggunakan library *sastrawi* untuk menghapus kata hubung (*stopword removal*), dan mengubah kata berimbuhan menjadi kata dasar (*stemming*). Lalu *tokenizer* untuk memecah kalimat menjadi kata tunggal.

ulasan	ulasan_case_folding	ulasan_cleaning	ulasan_normalisasi	ulasan_stopword	clean_ulasan	tokenizing
gak worth it	gak worth it	gak worth it	gak worth it	worth it	worth it	[worth, it]
pengalaman saya cukup baik rasanya tak ada masalah	pengalaman saya cukup baik rasanya tak ada masalah	pengalaman saya cukup baik rasanya tak ada masalah	pengalaman saya cukup baik rasanya tak ada masalah	alam cukup baik rasanya ada masalah info, akun sam...	alam cukup baik rasanya ada masalah info, akun sam...	[alam, cukup, baik, rasanya, ada, masalah, info, akun, sam]
lomakash 24 jam on	lomakash 24 jam on	lomakash 24 jam on	lomakash 24 jam on	terima kasih jam on	terima kasih jam on	[terima, kasih, jam, on]
mohon di kasih mode chat yg bisa kirim foto an...	mohon di kasih mode chat yg bisa kirim foto an...	mohon di kasih mode chat yg bisa kirim foto an...	mohon di kasih mode chat yg bisa kirim foto an...	mohon di kasih mode chat yg bisa kirim foto an...	mohon di kasih mode chat yg bisa kirim foto an...	[mohon, di, kasih, mode, chat, yg, bisa, kirim, foto, an]

Gambar 5. Hasil *preprocessing data*

4.3. Pembagian Data (Split Data)

Selanjutnya setelah *preprocessing data* maka dilakukan pembagian menggunakan menjadi 80:20 untuk data latih dan uji.

```

from sklearn.model_selection import train_test_split

df = pd.read_csv('data_maxim_preprocessed.csv', sep=';')
df['clean_ulasan'] = df['clean_ulasan'].fillna('')

X = df['clean_ulasan']
y = df['sentimen']

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

print(f"Jumlah Data Latih: {len(X_train)}")
print(f"Jumlah Data Uji: {len(X_test)}")

Jumlah Data Latih: 1932
Jumlah Data Uji: 484

```

Gambar 6. Pembagian data

4.4. Pembobotan Dengan TF-IDF

Setelah pembagian data selesai, selanjutnya tahap yang dilakukan adalah pembobotan nilai menggunakan TF-IDF dengan rumus dibawah ini

$$TF(t, d) = f_{t,d} \quad (2)$$

$TF(t, d)$: Nilai frekuensi term, seberapa kuat bobot sebuah *term* t di satu dokumen ulasan d

$f_{t,d}$: Jumlah kemunculan kata t pada ulasan d .

$$IDF(t) = \log\left(\frac{N}{df_t}\right) \quad (3)$$

$IDF(t)$: Jumlah IDF dari kata t .

N : Jumlah dari seluruh ulasan yang digunakan $N = 2.416$

df_t : jumlah ulasan yang memuat kata t

$$W_{t,d} = TF(t, d) \times IDF(t) \quad (4)$$

$W_{t,d}$: Bobot akhir kata t pada ulasan d .

```

tfidf = TfidfVectorizer()
X_train_tfidf = tfidf.fit_transform(X_train)
X_test_tfidf = tfidf.transform(X_test)

print("\nProses TF-IDF Selesai!")
print(f"Jumlah Fitur (Kata Unik): {X_train_tfidf.shape[1]}")

nama_fitur = tfidf.get_feature_names_out()
df_tfidf = pd.DataFrame(X_train_tfidf.toarray(), columns=nama_fitur)

print("Matriks TF-IDF:")
display(df_tfidf.iloc[:5, 100:110])

```

	asa	asik	asli	astaga	asyik	atas	atau	atm	atribut	attitude
0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Gambar 7. Pembobotan nilai TF-IDF

4.5. Penerapan Naive Bayes Classifier

Kemudian, *Naive Bayes Classifier* diterapkan guna menentukan kategori sentimen tersebut.

```

from sklearn.naive_bayes import MultinomialNB
from sklearn.metrics import accuracy_score, classification_report, confusion_matrix
import seaborn as sns
import matplotlib.pyplot as plt

model_nb = MultinomialNB()
model_nb.fit(X_train_tfidf, y_train)

y_pred = model_nb.predict(X_test_tfidf)

```

Gambar 8. Penerapan *Naive Bayes Classifier*

4.6. Evaluasi Model Klasifikasi

Tahap selanjutnya adalah melakukan evaluasi dengan *confusion matrix* guna mengukur akurasi, *precision*, *recall*, serta *f1-score*. Hasilnya adalah sebagai berikut:

```

accuracy = accuracy_score(y_test, y_pred)

print(f"--- HASIL EVALUASI MODEL ---")
print(f"Akurasi: {accuracy * 100:.2f}%")
print("\nLaporan Klasifikasi:")
print(classification_report(y_test, y_pred))

```

```

--- HASIL EVALUASI MODEL ---
Akurasi: 94.01%

Laporan Klasifikasi:

```

	precision	recall	f1-score	support
negatif	0.92	0.65	0.76	71
positif	0.94	0.99	0.97	413
accuracy			0.94	484
macro avg	0.93	0.82	0.86	484
weighted avg	0.94	0.94	0.94	484

Gambar 9. Hasil evaluasi model klasifikasi

Untuk penelitian berikutnya, disarankan untuk menggunakan metode penanganan ketidakseimbangan data (*data imbalance*), seperti SMOTE (*Synthetic Minority Over-sampling Technique*). Metode tersebut diperlukan agar meningkatkan nilai *recall* pada kelas sentimen negatif agar model lebih mengenali ulasan keluhan.

DAFTAR PUSTAKA

- [1] Saripah Aini Pohan, Samsudin, and Fathiya Hasyifah Sibarani, "Analisis Sentimen Terhadap Aplikasi Maxim Menggunakan Algoritma Random Forest," *J. Sci. Soc. Res.*, vol. 7, no. 3, pp. 1201–1208, 2024, doi: <https://doi.org/10.54314/jssr.v7i3.2151>.
- [2] Muhammad Nur Akbar, Nur Hasanahlmariyah Rusydi, M. Hasrul H., Nurul Shaumi Ramadhanti, and Erfiana, "Sentiment Analysis of Review Aplikasi Maxim di Google Play Store Menggunakan Support Vector Machine (SVM)," *AGENTS (Jurnal Sist. Informasi)*, vol. 2, no. 2, pp. 1–8, 2022, [Online]. Available: <https://doi.org/10.24252/jagti.v2i2.39>
- [3] O. I. Gifari, M. Adha, F. Freddy, and F. F. S. Durrand, "Analisis Sentimen Review Film Menggunakan TF-IDF dan Support Vector Machine," *J. Inf. Technol.*, vol. 2, no. 1, pp. 36–40, Mar. 2022, doi: [10.46229/jifotech.v2i1.330](https://doi.org/10.46229/jifotech.v2i1.330).
- [4] D. Chrisinta and J. E. Simarmata, "Eksplorasi Teknik Web Scraping pada Data Mining: Pendekatan Pencarian Data Berbasis Python," *Fakt. Exacta*, vol. 17, no. 1, pp. 58–68, May 2024, doi: [10.30998/faktorexacta.v17i1.22393](https://doi.org/10.30998/faktorexacta.v17i1.22393).
- [5] S. A. S. Mola, R. V. K. I. O. Roma, and T. Widiastuti, *Text Mining: Analisis Sentimen dengan Lexicon*. Kaizen Media Publishing, 2025. [Online]. Available: <https://books.google.co.id/books?id=sJ90EQAAQBAJ>
- [6] A. Septiani, A. Voutama, S. Siska, A. Andri Hendriadi, and N. Heryana, "ANALISIS SENTIMEN MENGGUNAKAN ALGORITMA NAÏVE BAYES TERHADAP REGULASI TIKTOK SHOP PADA MEDIA SOSIAL X (TWITTER)," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 8, no. 4, pp. 5729–5735, Jun. 2024, doi: [10.36040/jati.v8i4.10040](https://doi.org/10.36040/jati.v8i4.10040).
- [7] I. H. Kusuma and N. Cahyono, "Analisis Sentimen Masyarakat Terhadap Penggunaan E-Commerce Menggunakan Algoritma K-Nearest Neighbor," *J. Inform. J. Pengemb. IT*, vol. 8, no. 3, pp. 302–307, 2023, doi: [10.30591/jpit.v8i3.5734](https://doi.org/10.30591/jpit.v8i3.5734).
- [8] I. R. Hendrawan and E. Utami, *Natural Language Processing: Eksplorasi Sentimen Masyarakat dalam Evaluasi Produk Lokal Indonesia menggunakan Algoritma Bag of Words, TF-IDF, Word2Vec, dan Doc2Vec*. Penerbit Andi, 2023. [Online]. Available: <https://books.google.co.id/books?id=caPgEAAAQBAJ>
- [9] R. Saputra and F. N. Hasan, "Analisis Sentimen Terhadap Program Makan Siang & Susu Gratis Menggunakan Algoritma Naive Bayes," *J. Teknol. Dan Sist. Inf. Bisnis*, vol. 6, no. 3, pp. 411–419, 2024, doi: [10.47233/jteksis.v6i3.1378](https://doi.org/10.47233/jteksis.v6i3.1378).
- [10] A. Wibowo, Firman Noor Hasan, L. A. Ramadhan, Rika Nurhayati, and Arief Wibowo, "Analisis Sentimen Opini Masyarakat Terhadap Keefektifan Pembelajaran Daring Selama Pandemi COVID-19 Menggunakan Naïve Bayes Classifier," *J. Asimetrik J. Ilm. Rekayasa Inov.*, pp. 239–248, Jul. 2022, doi: [10.35814/asiimetrik.v4i1.3577](https://doi.org/10.35814/asiimetrik.v4i1.3577).
- [11] M. Agustriya, M. Ula, and K. -, "Analisis Kinerja Algoritma Klasifikasi Naïve Bayes Menggunakan Genetic Algorithm dan Bagging untuk Data Publik Risiko Transaksi Kartu Kredit," *J. Sist. dan Teknol. Inf.*, vol. 12, no. 3, p. 584, 2024, doi: [10.26418/justin.v12i3.80136](https://doi.org/10.26418/justin.v12i3.80136).
- [12] Suci Amaliah, M. Nusrang, and A. Aswi, "Penerapan Metode Random Forest Untuk Klasifikasi Varian Minuman Kopi di Kedai Kopi Konijiwa Bantaeng," *VARIANSI J. Stat. Its Appl. Teach. Res.*, vol. 4, no. 3, pp. 121–127, Dec. 2022, doi: [10.35580/variensiunm31](https://doi.org/10.35580/variensiunm31).
- [13] D. R. Alghifari, M. Edi, and L. Firmansyah, "Implementasi Bidirectional LSTM untuk Analisis Sentimen Terhadap Layanan Grab Indonesia," *J. Manaj. Inform.*, vol. 12, no. 2, pp. 89–99, Sep. 2022, doi: [10.34010/jamika.v12i2.7764](https://doi.org/10.34010/jamika.v12i2.7764).
- [14] M. Saifurridho, M. Martanto, and U. Hayati, "Analisis Algoritma K-Nearest Neighbor terhadap Sentimen Pengguna Aplikasi Shopee," *J. Inform. Terpadu*, vol. 10, no. 1, pp. 21–26, Mar. 2024, doi: [10.54914/jit.v10i1.1054](https://doi.org/10.54914/jit.v10i1.1054).
- [15] Pande sindu, Agus Aan Jiwa Permana, and I Nyoman Saputra Wahyu Wijaya, "IDENTIFIKASI DAN NORMALISASI TEKS SLANG DENGAN FASTTEXT PADA TWITTER DALAM BAHASA INDONESIA," *J. Pendidik. Teknol. dan Kejur.*, vol. 21, no. 1, pp. 33–44, Jan. 2024, doi: [10.23887/jptkundiksha.v21i1.66381](https://doi.org/10.23887/jptkundiksha.v21i1.66381).