

ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI GRAB MOBILE MENGUNAKAN METODE *K-NEAREST NEIGHBOR* DAN *LEXICON-BASED*

Maulana Bakti[✉], Pitrasacha Adytia, Bartolomius Harpad

Program Studi Sistem Informasi, STMIK Widya Cipta Dharma, Samarinda, Indonesia

Email: 2241026@wicida.ac.id

DOI: <https://doi.org/10.46880/jmika.Vol10No1.pp402-411>

ABSTRACT

Grab is an online transportation application and digital service that is widely used in Indonesia. User reviews on Google Play Store can serve as a source of information to understand user perceptions and satisfaction with the services provided. This study aims to analyze the sentiment of Grab application user reviews using a combination of the Lexicon-Based method for automatic labeling and the K-Nearest Neighbor (KNN) algorithm for classification. The dataset consists of 2.467 Indonesian-language reviews that undergo preprocessing, lexicon-based sentiment labeling, and feature extraction using Term Frequency-Inverse Document Frequency (TF-IDF). The classification process is performed using KNN with $K=5$ and an 80:20 data split ratio. The results show that 80% of the reviews contain positive sentiment. The KNN model achieves an accuracy of 92,71%, precision of 92.48%, recall of 98.08%, and F1-score of 95.20%. This study demonstrates that the combination of Lexicon-Based and KNN methods is effective for sentiment analysis of Indonesian-language reviews and can support service quality evaluation based on user opinions.

Keyword: *Sentiment Analysis, K-Nearest Neighbor, Lexicon-Based, TF-IDF, Indonesian Language.*

ABSTRAK

Grab merupakan aplikasi transportasi online dan layanan digital yang banyak digunakan di Indonesia. Ulasan pengguna pada Google Play Store dapat menjadi sumber informasi untuk mengetahui persepsi dan kepuasan pengguna terhadap layanan. Penelitian ini bertujuan untuk menganalisis sentimen ulasan pengguna aplikasi Grab menggunakan kombinasi metode Lexicon-Based untuk pelabelan otomatis dan algoritma K-Nearest Neighbor (KNN) untuk klasifikasi. Dataset penelitian terdiri dari 2.467 ulasan berbahasa Indonesia yang melalui tahapan preprocessing, pelabelan sentimen berbasis leksikon, serta ekstraksi fitur menggunakan Term Frequency-Inverse Document Frequency (TF-IDF). Proses klasifikasi dilakukan menggunakan KNN dengan nilai $K=5$ dan pembagian data 80:20. Hasil penelitian menunjukkan bahwa 80% ulasan memiliki sentimen positif. Model KNN menghasilkan accuracy sebesar 92,71%, precision sebesar 92,48%, recall sebesar 98,08%, dan F1-score sebesar 95,20%. Penelitian ini menunjukkan bahwa kombinasi Lexicon-Based dan KNN efektif untuk analisis sentimen ulasan berbahasa Indonesia serta dapat mendukung evaluasi kualitas layanan berdasarkan opini pengguna.

Kata Kunci: *Analisis Sentimen, K-Nearest Neighbor, Lexicon-Based, TF-IDF, Bahasa Indonesia.*

PENDAHULUAN

Kemajuan pesat dalam dunia teknologi informasi dan komunikasi telah menjadi faktor pendorong tumbuhnya aplikasi berbasis *mobile* di beragam sektor kehidupan, termasuk di dalamnya layanan transportasi online. Grab merupakan salah satu platform transportasi online dan layanan digital terbesar di Asia Tenggara yang digunakan secara luas oleh masyarakat Indonesia. Layanan yang disediakan Grab mencakup transportasi, pengiriman makanan, belanja online, hingga layanan keuangan digital (Ridwan Usmana & Rais Ruli, 2025).

Seiring meningkatnya jumlah pengguna, platform *Google Play Store* menjadi wadah bagi jutaan pengguna untuk mengekspresikan pengalaman mereka

melalui ulasan dan rating. Ulasan tersebut mengandung informasi berharga terkait kualitas layanan, fitur aplikasi, harga, kinerja driver, dan performa sistem secara keseluruhan. Namun demikian, volume ulasan yang sangat besar membuat analisis manual sulit (Sari et al., 2017).

Salah satu cabang proses pengolahan bahasa *Natural Language Processing (NLP)* adalah analisis sentimen, yang bertujuan untuk menentukan orientasi opini dari suatu teks, apakah bernilai positif, negatif, atau netral (Dyo fatra et al., 2020). Teknologi ini dapat diterapkan untuk mengekstrak wawasan dari ulasan pengguna secara otomatis dan produktif.

Metode yang digunakan dalam penelitian ini *KNN* sebagai algoritma klasifikasi dikombinasikan

dengan pendekatan *Lexicon-Based* untuk pelabelan sentimen secara otomatis. *KNN* merupakan algoritma pembelajaran mesin yang sederhana namun efektif, yang mengklasifikasikan data baru berdasarkan kedekatan dengan data latih yang sudah ada (Legito et al., 2023). Sementara itu, pendekatan *Lexicon-Based* menggunakan kamus sentimen untuk memberikan skor pada teks sehingga dapat dilabeli secara otomatis tanpa memerlukan pelabelan manual (Taboada et al., 2011).

Beberapa penelitian sebelumnya telah mengkaji analisis sentimen pada berbagai platform ulasan digital (Riansyah Ramadhan & Agus Sugianto, 2024). Melakukan analisis sentimen ulasan aplikasi Gojek menggunakan *Naive Bayes* dan mendapatkan akurasi sebesar 84,2%. Menggunakan *KNN* untuk klasifikasi sentimen ulasan *e-commerce* dengan akurasi 81,5% (Syahputri et al., 2024). Menerapkan pendekatan *Lexicon-Based* pada ulasan hotel berbahasa Indonesia dan berhasil mencapai *F1-score* rata-rata sebesar 0,78. Penelitian ini memperbarui dan mengembangkan pendekatan tersebut dengan menggabungkan *Lexicon-Based* dan *KNN* pada dataset ulasan aplikasi Grab.

Namun, beberapa penelitian sebelumnya masih memiliki keterbatasan, seperti penggunaan metode klasifikasi tanpa dikombinasikan dengan teknik pelabelan otomatis, sehingga proses pelabelan data masih banyak dilakukan secara manual dan memakan waktu yang tidak sedikit. Selain itu, penelitian terkait analisis sentimen pada ulasan aplikasi transportasi online khususnya aplikasi Grab masih relatif terbatas, terutama yang menerapkan kombinasi metode *Lexicon-Based* dan algoritma *KNN*. Oleh karena itu, penelitian ini mengusulkan kombinasi metode *Lexicon-Based* sebagai pelabelan sentimen otomatis dan *KNN* sebagai metode klasifikasi untuk menganalisis sentimen ulasan pengguna aplikasi Grab pada *Google Play Store*.

Meskipun pendekatan *Lexicon-Based* dan algoritma *KNN* telah banyak digunakan dalam penelitian analisis sentimen, penelitian ini memberikan kontribusi melalui penerapan kamus sentimen bahasa Indonesia (InSet) sebagai mekanisme pelabelan sentimen otomatis sebelum proses klasifikasi menggunakan *KNN* pada ulasan pengguna aplikasi Grab di *Google Play Store*. Berbeda dengan beberapa penelitian terdahulu yang menggunakan pendekatan klasifikasi atau strategi pelabelan yang berbeda, penelitian ini mengintegrasikan pelabelan otomatis, preprocessing, ekstraksi fitur *TF-IDF*, dan klasifikasi *KNN* dalam satu alur analisis sehingga mampu mengurangi kebutuhan pelabelan manual pada data ulasan berbahasa Indonesia. Hasil penelitian selanjutnya dianalisis berdasarkan metrik accuracy,

precision, recall, dan *F1-score* untuk mengevaluasi performa pendekatan yang diterapkan pada domain analisis sentimen ulasan aplikasi transportasi online.

Dari latar belakang yang ditunjukkan, persoalan pertama yang diangkat dalam penelitian ini adalah, Seperti apa persebaran sentimen pengguna terhadap aplikasi Grab berdasarkan komentar dan penilaian yang tersedia di *Google Play Store*? Bagaimana penerapan metode *Lexicon-Based* dalam pelabelan sentimen ulasan berbahasa Indonesia? Bagaimana performa metode *KNN* dalam mengklasifikasikan sentimen untuk ulasan pengguna aplikasi Grab berdasarkan nilai akurasi, presisi, *recall*, dan *F1-score*.

Tujuan dari penelitian ini adalah untuk, mengetahui distribusi sentimen pengguna untuk aplikasi Grab berdasarkan data ulasan yang terdapat di *Google Play Store*, mengkategorikan ulasan pengguna menjadi kategori sentimen positif dan negatif. menggunakan *Lexicon-Based*, mengukur performa metode *KNN* dalam klasifikasi sentimen ulasan berbahasa Indonesia serta, menerapkan *pipeline NLP* lengkap meliputi *preprocessing*, *TF-IDF*, dan *KNN* untuk analisis sentimen.

KAJIAN LITERATUR

Analisis sentimen termasuk dalam salah satu bidang kajian (*NLP*) yang memiliki tujuan untuk mendeteksi dan mengungkap perasaan atau pandangan seseorang yang tercermin dalam sebuah teks, apakah cenderung bersifat positif atau negatif dalam menganalisis tanggapan pengguna aplikasi digital, analisis sentimen berperan penting sebagai alat bantu untuk memahami persepsi masyarakat secara otomatis dari data teks berskala besar. Ridwan Usman dan Rais Ruli (2025) dalam studi tentang analisis sentimen ulasan pengguna aplikasi Kereta Api Indonesia (KAI) *Access*, ditemukan bahwa komentar oleh pengguna di platform seperti *Google Play Store* menyimpan informasi yang sangat bernilai guna mengevaluasi kualitas layanan transportasi online secara objektif. Penelitian tersebut menggunakan pendekatan *Deep Learning* dengan algoritma *Long Short-Term Memory (LSTM)*, *Gated Recurrent Unit (GRU)*, dan *Convolutional Neural Network CNN* untuk mengolah teks ulasan berbahasa Indonesia (Riansyah Ramadhan & Agus Sugianto, 2024), dan membuktikan bahwa pemrosesan teks secara otomatis mampu memberikan gambaran yang komprehensif terhadap kepuasan pengguna layanan transportasi berbasis aplikasi. Hasil dari penelitian ini menjadi salah satu dasar pemikiran bahwa analisis sentimen berbasis ulasan aplikasi transportasi online merupakan pendekatan yang valid dan relevan dalam mengukur persepsi pengguna secara

kuantitatif. Penelitian ini merupakan penelitian terapan dengan pendekatan kuantitatif yang berfokus pada penerapan metode machine learning dan *NLP* untuk analisis sentimen teks berbahasa Indonesia. Alur penelitian terdiri dari beberapa tahapan utama.

Lexicon-Based adalah metode analisis sentimen yang bekerja dengan mencocokkan kata-kata yang terdapat dalam teks yang memiliki kamus sentimen (*sentiment lexicon*) yang mencakup daftar kata beserta polaritas sentimennya, baik positif atau negatif. Skor sentimen suatu dokumen dihitung berdasarkan perbedaan antara banyaknya kata positif dan kata negatif yang ditemukan dalam teks. Jika skor yang diproduksi lebih besar atau sama dengan nol, maka Teks dianggap sebagai sentimen positif. Sebaliknya jika skor negatif, Teks dianggap sebagai sentimen negatif. Dyo Fatra, Hayatin, dan Aditya (2020) dalam penelitiannya yang berjudul "Analisa Sentimen *Tweet* Berbahasa Indonesia dengan Menggunakan Metode *Lexicon* pada Topik Perpindahan Ibu Kota Indonesia" menjelaskan bahwa metode *Lexicon-Based* merupakan pendekatan yang efektif untuk menganalisis sentimen teks berbahasa Indonesia tanpa memerlukan proses pelatihan model secara eksplisit. Penelitian tersebut menunjukkan bahwa dengan memanfaatkan kamus sentimen yang sesuai dengan karakteristik bahasa target, dengan metode ini, proses identifikasi sentimen dapat dilakukan secara otomatis dan menghasilkan performa yang cukup baik, terutama dalam hal analisis data teks yang tidak formal seperti postingan pada platform media sosial. Keunggulan utama metode ini terletak pada kesederhanaan implementasinya serta kemampuannya untuk digunakan sebagai mekanisme pelabelan otomatis (*automatic labeling*) pada dataset yang belum memiliki label sentimen sebagaimana diterapkan dalam penelitian ini.

KNN adalah algoritma *machine learning* berbasis instance yang mengklasifikasikan suatu data baru yang diperoleh berdasarkan kebanyakan label dari K data tetangga terdekat di area fitur. Jarak antar data dihitung menggunakan formula tertentu, yang paling umum adalah *Euclidean Distance*. Algoritma *KNN* tergolong sederhana namun terbukti efektif.

Untuk beragam tugas klasifikasi teks, salah satunya termasuk analisis sentimen. (Legito et al., 2023) dalam penelitiannya yang berjudul "Aplikasi algoritma *KNN* untuk analisis sentimen tentang radikalisme dan unsur kekhalfahan di Indonesia" mendemonstrasikan secara empiris bahwa algoritma *KNN* memiliki kemampuan untuk mengkategorikan teks sentimen berbahasa Indonesia dengan performa yang memadai (Syahputri et al., 2024). Penelitian tersebut membuktikan bahwa *KNN* dapat diterapkan

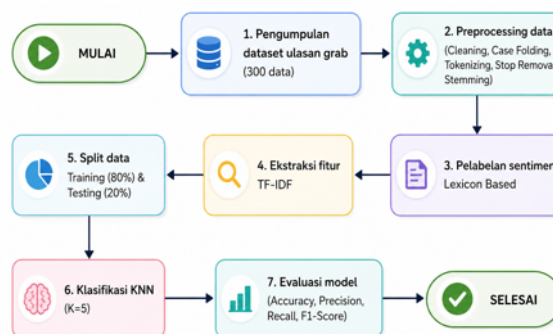
secara efektif pada data teks yang diwakili oleh vektor fitur numerik, dan menghasilkan akurasi klasifikasi yang kompetitif dibandingkan dengan algoritma lainnya. Temuan ini memperkuat dasar pemilihan algoritma *KNN* dalam penelitian ini sebagai classifier utama yang dikombinasikan dengan representasi fitur *TF-IDF* dan pelabelan otomatis berbasis *Lexicon-Based* untuk menganalisis sentimen ulasan pengguna aplikasi Grab.

METODE PENELITIAN

Alur Penelitian

Penelitian ini terapan dengan metode kuantitatif yang berfokus pada penerapan metode *machine learning* dan *NLP* untuk menganalisa sentimen teks berbahasa Indonesia (Tribuana et al., 2025). Penelitian dilakukan melalui beberapa tahapan yang saling berkaitan, mulai dari pengumpulan dataset, preprocessing data, pelabelan sentimen, ekstraksi fitur, proses klasifikasi, hingga evaluasi model untuk memperoleh hasil dari analisis sentimen yang akurat berdasarkan ulasan pengguna aplikasi Grab.

Alur penelitian terdiri dari beberapa tahapan utama yang ditunjukkan pada gambar 1.



Gambar 1. Metode Penelitian

Berdasarkan Gambar 1, penelitian dimulai dengan proses pengumpulan dataset berupa Pengumpulan dataset ulasan grab 2.467 data Grab dari Google Play Store. Dataset yang telah dikumpulkan selanjutnya dilakukan tahapan preprocessing yang meliputi *cleaning*, *case folding*, *tokenizing*, *stopword removal*, dan *stemming* untuk membersihkan serta menormalisasi teks. Setelah proses preprocessing selesai, data diberikan label sentimen menggunakan metode *Lexicon-Based* untuk menentukan kategori sentimen positif dan negatif secara otomatis. Selanjutnya dilakukan ekstraksi fitur menggunakan metode *TF-IDF* untuk mengubah data teks menjadi representasi numerik yang dapat diproses oleh algoritma klasifikasi. Data kemudian dibagi menjadi data training sebesar 80% dan data testing sebesar 20%.

Tahap berikutnya adalah proses klasifikasi menggunakan algoritma KNN dengan nilai $K=5$. Hasil klasifikasi kemudian dievaluasi menggunakan confusion matrix dengan parameter accuracy, precision, recall, dan F1-score untuk mengetahui performa model dalam mengklasifikasikan sentimen ulasan pengguna aplikasi Grab.

Dataset

Ulasan pengguna adalah dataset yang digunakan dalam penelitian ini. aplikasi Grab yang dikumpulkan dari Google Play Store. Total dataset berjumlah 2.467 ulasan berbahasa Indonesia. Setiap data terdiri dari empat atribut, yaitu: nomor urut (No), teks ulasan (Ulasan), rating bintang (Rating), dan nama pengguna (Nama). Pengumpulan Data dilakukan secara manual. Pengambilan data dari platform Google Play Store.

Preprocessing Data

Tahap *preprocessing* merupakan langkah krusial untuk menjamin kualitas informasi sebelum diolah oleh model. Proses *preprocessing* yang diterapkan meliputi lima tahapan (Setyani et al., 2026). (1) *Cleaning Text* teks ulasan dibersihkan dari unsur-unsur yang tidak penting seperti *URL*, angka, tanda baca, dan spasi berlebih, (Rifaldi & Fadlil, 2023) "bagus!!! 😊 #recommended" menjadi "bagus recommended". (2) *Case Folding* mengubah huruf besar atau kecil, diubah menjadi huruf kecil semua, "Driver Ramah Sekali" menjadi "driver ramah sekali". (3) *Tokenizing* memisahkan teks menjadi kata menggunakan fungsi *word_tokenize* dari pustaka *NLTK*, (Rifaldi & Fadlil, 2023) "driver ramah sekali" menjadi "driver", "ramah", "sekali". (4) *Stopword Removal* penghapusan kata-kata umum yang tidak signifikan, (Rifaldi & Fadlil, 2023) "driver", "ramah", "sekali" menjadi "driver", "ramah". (5) *Stemming* perubahan kata ke bentuk dasarnya menggunakan pustaka *PySastrawi*, (Pardede, 2025) "keterlambatan" menjadi "lambat".

Pelabelan Sentimen (Lexicon-Based)

Metode berbasis *lexicon* digunakan untuk melabelkan sentimen secara otomatis. menggunakan kamus sentimen Indonesia *Sentiment lexicon* (InSet). Kamus sentimen yang digunakan mencakup kata positif (bagus, baik, cepat, mantap, murah, nyaman, ramah, hebat, puas, suka) dan kata negatif (jelek, buruk, lama, mahal, kecewa, error, parah, lemot, gagal, batal). Skor sentimen dihitung menggunakan persamaan (Amal & Jayanta, 2023).

$$Score(d) = \frac{\sum pos(w) - \sum neg(w)}{\sum neg(w)} \quad (1)$$

Jika $Score(d) \geq 0$, ulasan diberi label positif; jika $Score(d) < 0$, ulasan diberi label negatif. Penentuan label ini didasarkan pada hasil perhitungan skor sentimen dari setiap ulasan, di mana nilai skor mencerminkan kecenderungan opini yang terkandung dalam teks. Tabel 3 menampilkan contoh 10 ulasan beserta hasil pelabelan menggunakan data dari kamus InSet dengan kamus positif sebanyak 3.609 kata dan kamus negatif sebanyak 6.609 kata.

Ekstraksi Fitur TF-IDF

TF-IDF (*Term Frequency-Inverse Document Frequency*) digunakan untuk merepresentasikan setiap dokumen sebagai vektor numerik. (Maheri et al., 2025) sehingga dapat memudahkan proses analisis oleh algoritma klasifikasi. Metode ini menggabungkan dua elemen penting, *Term Frequency* (TF), yang menentukan seberapa sering kata ditampilkan dalam dokumen, serta *Inverse Document Frequency* (IDF) yang mengukur tingkat pentingnya sebuah kata tersebut berdasarkan kelangkaannya pada seluruh dokumen. Semakin kata sering muncul dalam dokumen, dan jumlah mereka meningkat jarang muncul, maka bobot kata tersebut akan semakin besar. Rumus yang digunakan adalah.

$$IDF(t) = \log_{10} \left(\frac{N}{df(t)} \right) \quad (2)$$

$$TF - IDF(t, d) = TF(t, d) \times IDF(t) \quad (3)$$

N = jumlah total dokumen (2.467) dan $df(t)$ = jumlah dokumen yang mengandung kata t .

Klasifikasi K-Nearest Neighbor

Algoritma *KNN* merupakan metode klasifikasi berbasis kemiripan yang mengidentifikasi kelas data baru dengan melihat kedekatannya terhadap data lain yang sudah diketahui labelnya. Algoritma ini bekerja dengan mencari sejumlah K tetangga terdekat, kemudian menentukan kelas berdasarkan mayoritas label dari tetangga tersebut. Algoritma *KNN* mengklasifikasikan dokumen baru berdasarkan mayoritas label dari K tetangga terdekat. Dalam penelitian ini digunakan $K=5$ dengan metode voting mayoritas. Jarak antar dokumen dihitung menggunakan Euclidean Distance (Maldini & Septi, 2025).

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (4)$$

Pembagian Data dan Evaluasi Model

Dataset dibagi dengan rasio 80:20 menggunakan metode *random sampling* dengan

random_state=42 (Angelina et al., 2023). Evaluasi dilakukan menggunakan *Confusion Matrix* pada 494 data uji evaluasi model berasal dari kumpulan ulasan pengguna yang diperoleh dari aplikasi Grab dari *Google Play Store*, dengan total 2.467 ulasan berbahasa Indonesia. Dataset dibagi menggunakan rasio 80:20 dengan metode *random sampling* random_state=42, menghasilkan 1.973 ulasan data latih, dan 494 ulasan data uji 364 positif dan 130 negatif. Rumus yang digunakan adalah.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (5)$$

$$= 6 \text{ (Positif)}$$

HASIL DAN PEMBAHASAN

Hasil Pembahasan *Lexicon-Based*

Setiap kata yang termasuk dalam kamus positif akan memberikan nilai tambah, sedangkan kata dalam kamus negatif akan mengurangi nilai skor. Dengan demikian, nilai akhir dari perhitungan ini akan menunjukkan kecenderungan sentimen suatu ulasan, apakah bersifat positif atau negatif berdasarkan pada tabel 1.

$$Score (d) = \sum pos (w) - \sum neg(w)$$

$$= (3 + 3) - (0)$$

$$= 6 - 0$$

Tabel 1. Hasil Pelabelan Sentimen Menggunakan *Lexicon-Based*

No	Ulasan Asli	Pre processing	Kata Positif	Kata Negatif	Skor	Label
1	bagus dan membantu sekali	bagus bantu	bagus,	-	+3	Positif
2	pelayanan bagus menurut saya oke	layan bagus menurut oke	bagus, oke	-	+6	Positif
3	bagus	bagus	bagus	-	+3	Positif
4	saya tidak bisa login akibat capture muka tidak ada	login akibat capture muka berdelay gagal	-	gagal	-5	Negatif
5	cepat, aman, menyenangkan pengemudi komunikatif	cepat aman senang kemudi komunikatif	cepat, aman, senang	-	+9	Positif
6	sangat baik dan ramah drivernya	ramah drivernya	ramah	-	+3	Positif
7	bagus pilihanmu yang tepat	bagus pilih	bagus	-	+3	Positif
8	cepat cara ordernya	cepat order nya	cepat	-	+3	Positif
9	sangat aman dan nyaman	aman nyaman	aman, nyaman	-	+6	Positif
10	tadinya saya suka pakai jasa grab karena banyak promo	suka pakai jasa grab promonyadan gegara seliweran	suka, promo, promo	-	+9	Positif

Berdasarkan Tabel 1, pelabelan sentimen dilakukan dengan mengidentifikasi katattpositif dan negatif pada setiap ulasan. Skor dihitung dari selisih jumlah kata positif dan negatif, yang kemudian menentukan label akhir. Skor positif menunjukkan sentimen positif, sedangkan skor negatif menunjukkan sentimen negatif, sehingga metode *Lexicon-Based* dapat memberi pelabelan secara sistematis.(Purnamasari et al., 2023)

Hasil Perhitungan TF-IDF

Perhitungan *TF-IDF* dilakukan untuk mengetahui bobot kata-kata yang penting dalam dokumen terhadap keseluruhan dataset. Pada tahap ini,

terlebih dahulu dihitung nilai *IDF* untuk mengukur tingkat kelangkaan suatu kata, kemudian dikalikan dengan nilai *TF* untuk memperoleh bobot akhir kata tersebut dalam dokumen.

$$IDF(t) = \log_{10} \left(\frac{2467}{75} \right)$$

$$= \log_{10} (32,8933)$$

$$= 1,571$$

Setelah mendapatkan hasil dari perhitungan *IDF* langkah selanjutnya mencari hasil dari perhitungan *TF-IDF*.

$$TF-IDF = 1 \times 32,8933 = 1,571$$

Setelah mendapatkan hasil *TF-IDF* hasil perhitungan ini menunjukkan bahwa kata tersebut

memiliki tingkat kepentingan tertentu dalam dokumen berdasarkan frekuensi kemunculan dan kelangkaannya. Setelah mendapatkan hasil *TF-IDF*, pada tabel 2

ditampilkan contoh perhitungan *TF-IDF* untuk 9 kata representatif dari dataset.

Tabel 2. Perhitungan TF-IDF

Kata	Dok	TF	df	N	IDF = log(N/df)	TF-IDF
bantu	d1	1	75	2467	$\log(2467/75) = 1,571$	$1 \times 1,571 = 1,571$
bagus	d1	1	466	2467	$\log(2467/466) = 0,724$	$1 \times 0,724 = 0,724$
layan	d2	1	146	2467	$\log(2467/146) = 1,227$	$1 \times 1,227 = 1,227$
Oke	d3	1	64	2467	$\log(2467/64) = 1,568$	$1 \times 1,568 = 1,568$
berdelay	d3	1	1	2467	$\log(2467/1) = 3,392$	$1 \times 3,392 = 3,392$
login	d4	2	9	2467	$\log(2467/9) = 2,438$	$2 \times 2,438 = 4,876$
gagal	d4	1	10	2467	$\log(2467/10) = 2,392$	$1 \times 2,392 = 2,392$
komunikatif	d5	1	3	2467	$\log(2467/3) = 2,915$	$1 \times 2,915 = 2,915$
senang	d5	1	33	2467	$\log(2467/33) = 1,874$	$1 \times 1,874 = 1,874$

Dari hasil tabel 2 perhitungan *TF-IDF* setiap dokumen disusun ke dalam matriks vektor. Pada tabel

3 menampilkan representasi vektor *TF-IDF* dari 5 dokumen latihan dan 1 dokumen.

Tabel 3. Matrix Vektor TF-IDF

Dokumen	bantu	bagus	layan	berdelay	login	senang	Label
d1	1,571	0,724	0	0	0	0	Positif
d2	0	0	1,227	0	0	0	Positif
d3	0	0	0	3,392	0	0	Negatif
d4	0	0	0	0	4,876	0	Positif
d5	0	0	0	0	0	1,874	Positif
q (uji)	1,571	0	0	0	2,438	0	?

Perhitungan pada tabel 3 diatas jarak antara data uji dan data latihan dilakukan menggunakan metode Euclidean Distance. Data uji direpresentasikan sebagai vektor *TF-IDF*, kemudian dibandingkan dengan vektor *TF-IDF* pada data latihan. Jarak dihitung dengan mengurangi setiap atribut yang bersesuaian, mengkuadratkan hasilnya, menjumlahkannya, lalu diakarkan sehingga menghasilkan nilai jarak.

bagus puas" terhadap 5 dokumen latihan terdekat. Berdasarkan K=5, diperoleh 3 label Positif dan 2 label Negatif sehingga voting mayoritas menghasilkan prediksi Positif berdasarkan pada tabel 4.

Hasil K-Nearest Neighbor

Pada tabel 4 menampilkan contoh perhitungan *Euclidean Distance* untuk dokumen uji q = "aplikasi

$$\begin{aligned}
 d(q, d1) &= \sqrt{(1,571 - 1,571)^2 + (0 - 0,724)^2 + (2,438 - 0)^2} \\
 &= \sqrt{(0)^2 + (0,724)^2 + (2,438)^2} \\
 &= \sqrt{0,524 + 5,944} \\
 &= \sqrt{6,468} = 2,543
 \end{aligned}$$

Tabel 4. Perhitungan Euclidean Distance KNN (K=5)

Tetangga	Dok	Perhitungan Euclidean Distance	Jarak	Label
K-1	d1	$\sqrt{((1,571-1,571)^2 + (0-0,724)^2 + (0-2,438)^2)}$	2,543	Positif
K-2	d4	$\sqrt{((1,571)^2 + (2,438 - 4,876)^2)}$	2,900	Positif
K-3	d2	$\sqrt{((1,571-0)^2 + (0-1,227)^2 + (2,438-0)^2)}$	3,149	Positif
K-4	d5	$\sqrt{((1,571,-0)^2 + (2,438-0)^2 + (0-1,874)^2)}$	3,453	Positif
K-5	d3	$\sqrt{((1,571-0)^2 + (0-3,392)^2 + (2,438-0)^2)}$	4,463	Negatif

Berdasarkan tabel 4 K=5 tetangga terdekat diperoleh: 4 label Positif (d1, d4, d2, d5) dan 1 label Negatif (d3). Voting mayoritas menghasilkan prediksi label Positif untuk dokumen uji q. Hasil ini sesuai dengan label aktual, sehingga dikategorikan sebagai *True Positive (TP)*.

Pengujian Variasi Nilai K

Untuk menentukan nilai parameter K yang optimal pada algoritma *KNN*, penelitian ini melakukan pengujian terhadap beberapa variasi nilai K, yaitu K=3, K=5, dan K=7.

Tabel 5. Hasil Pengujian Variasi Nilai K

Nilai K	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
K=3	92.51	92.24	98.08	95,07
K=5	92.71	92.49	98.08	95.20
K=7	91.09	91.02	97.52	94.16

Hasil pengujian menunjukkan bahwa nilai K=5 memberikan performa terbaik dengan accuracy sebesar 93,72%, precision sebesar 92,49%, recall sebesar 98,08%, dan F1-score sebesar 92,20%. Meskipun nilai K=3 menghasilkan recall yang sama sebesar 98,08%, tetapi nilai accuracy, precision, dan F1-score mengalami penurunan dibandingkan dengan K=5. Oleh karena itu, nilai K=5 dipilih sebagai parameter yang digunakan pada model *KNN* karena memberikan keseimbangan performa klasifikasi yang lebih baik berdasarkan keseluruhan metrik evaluasi.

Evaluasi Model

Data evaluasi model berasal dari kumpulan ulasan pengguna yang diperoleh dari aplikasi Grab dari *Google Play Store*, dengan total 2.467 ulasan berbahasa Indonesia. Dataset dibagi menggunakan rasio 80:20 dengan metode *random sampling* (*random_state=42*), sehingga menghasilkan 1973 data latih dan 494 data uji yang terdiri dari 364 ulasan positif dan 130 ulasan negatif yang ditampilkan pada tabel 5.

Tabel 6. Confusion Matrix Hasil Pengujian

	Prediksi Positif	Prediksi Negatif	Total Aktual
Aktual Positif	TP = 357	FN = 7	364
Aktual Negatif	FP = 29	TN = 101	130
Total Prediksi	364	130	494

Dari data tabel 5 diatas menjelaskan aasil evaluasi menunjukkan bahwa *True Positif (TP)* = 357 ulasan berlabel prediksi yang berhasil positif secara tepat sebagai positif, *True Negatif (TN)* = 101 ulasan berlabel negatif yang diprediksi dengan benar secara tepat sebagai negatif, *False Positif (FP)* = sejumlah ulasan yang sesungguhnya negatif namun keliru diprediksi sebagai positif, serta *False Negatif (FN)* = sejumlah ulasan yang sesungguhnya positif namun keliru diprediksi sebagai negatif.

Dari nilai *Confusion Matrix* tersebut, diperoleh metrik evaluasi yang ditampilkan pada tabel 6. Hasil *matrix* evaluasi.

Perhitungan *matrix* evaluasi:

$$Accuracy = \frac{357+101}{357+101+29+7} = \frac{458}{494} = 92,71\%$$

$$Precision = \frac{TP}{TP+FP} = \frac{357}{357+29} = \frac{357}{386} = 92,49\%$$

$$Recall = \frac{TP}{TP+FN} = \frac{357}{357+7} = \frac{357}{364} = 98,08\%$$

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision+Recall} = 2 \times \frac{0,9249 \times 0,9808}{0,9249+0,9208} = 95,20\%$$

Tabel 7. Hasil Matrix Evaluasi

Metrik	Rumus	Perhitungan	Hasil
Accuracy	$\frac{(TP+TN)}{(TP+TN+FP+FN)}$	$\frac{(357+101)}{(357+101+29+7)}$ = 458/494	92,71%
Precision	$\frac{TP}{(TP+FP)}$	$\frac{357}{(357+29)}$ = 357/386	92,49%
Recall	$\frac{TP}{(TP+FN)}$	$\frac{357}{(357+7)}$ = 357/364	98,08%
F1-Score	$2 \times \frac{(P \times R)}{(P+R)}$	$\frac{2 \times (0,9249 \times 0,9208)}{(0,9249+0,9208)}$	95,20%

Berdasarkan temuan, pengujian menggunakan *Confusion Matrix*, berdasarkan hasil evaluasi, model menghasilkan nilai *accuracy* 92,71%, sedangkan *precision* 92,49%, *recall* 98,08%, serta *F1-score* sebesar 95,20%. Nilai *accuracy* bukti kemampuan model untuk mengklasifikasikan data dengan tingkat ketepatan yang tinggi.

Distribusi Sentimen

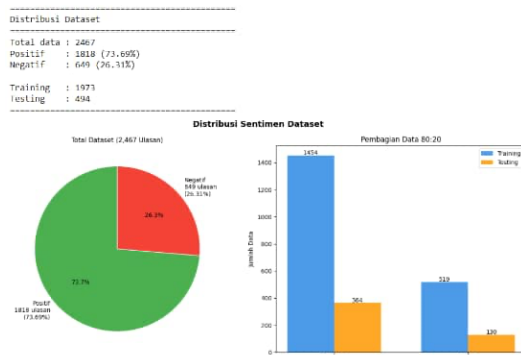
Setelah proses pelabelan menggunakan metode *Lexicon-Based* terhadap 2.467 ulasan, distribusi sentimen ditunjukkan pada Tabel 7. Sebanyak 1.973 ulasan (80%) dikategorikan sebagai sentimen positif

dan 494 ulasan (20%) dikategorikan sebagai sentimen negatif.

Tabel 8. Distribusi Sentimen Ulasan Pengguna Grab

Kelas Sentimen	Jumlah Ulasan	Persentase (%)
Positif	1.973	80%
Negatif	494	20%
Total	2.467	100%

Distribusi sentimen dan pembagian data training-testing secara visual ditunjukkan pada Gambar 2. Grafik tersebut memperjelas ketidak seimbangan distribusi sentimen antara kelas positif (80%) dan negatif (20%), serta proporsi data latih dan uji yang digunakan.



Gambar 2. Distribusi Sentimen Dataset dan Pembagian Data 80:20

Aspek Sentimen Layanan Grab

Selain mengetahui distribusi sentimen secara keseluruhan, penelitian ini juga menganalisis sentimen berdasarkan aspek layanan yang paling sering dibahas oleh pengguna aplikasi Grab. Analisis aspek dilakukan dengan mengelompokkan ulasan berdasarkan kemunculan kata kunci yang merepresentasikan fitur layanan tertentu, yaitu aspek driver, sapek aplikasi, dan aspek pelayanan. Pendekatan ini bertujuan untuk memberikan gambaran yang lebih rinci mengenai bagian layanan yang memperoleh apresiasi maupun keluhan dari pengguna.

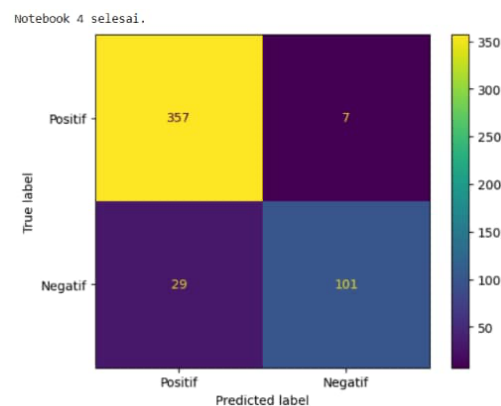
Analisis Confusion Matrix

Confusion Matrix pada data terfilter menunjukkan $TP = 357$, $TN = 103$, $FP = 29$, $FN = 7$ yang ditampilkan pada tabel 9 menyajikan *Confusion Matrix*.

Tabel 9. Confusion Matrix Data Terfilter Positif & Negatif

	Prediksi: Positif	Prediksi: Negatif	Total Aktual
Aktual: Positif	$TP = 357$	$FN = 7$	364
Aktual: Negatif	$FP = 29$	$TN = 101$	130
Total Prediksi	364	130	494

Berdasarkan Tabel 9, diperoleh dari total 494 data uji, sebanyak 458 data berhasil diklasifikasikan dengan benar dan 36 data mengalami kesalahan. Nilai TP dan TN yang tinggi menunjukkan bahwa model mampu untuk mengklasifikasikan data. dengan baik, sedangkan nilai FP dan FN yang rendah menunjukkan tingkat kesalahan yang kecil berikut yang di ditampilkan pada gambar 3.



Gambar 3. Confusion Matrix

Terlihat bahwa nilai terbesar berada pada bagian TP (357) dan TN (101), yang mengindikasikan bahwa sebagian besar data berhasil dikelompokkan ke dalam kelas yang benar. Sementara itu, nilai *False Positive* dan *False Negative* masing-masing sebesar 3 menunjukkan adanya kesalahan klasifikasi, namun jumlahnya relatif kecil dibandingkan dengan jumlah prediksi yang benar.

Pembahasan

Hasil penelitian menunjukkan bahwa penerapan metode (*KNN*) dengan nilai $K=5$ yang dikombinasikan dengan representasi fitur *TF-IDF* serta pelabelan otomatis menggunakan pendekatan *Lexicon-Based* mampu mengklasifikasikan sentimen ulasan pengguna aplikasi Grab dengan performa yang baik. Hal ini dibuktikan dengan nilai akurasi sebesar 92,71% pada data uji, yang menunjukkan bahwa model mampu mengenali pola sentimen secara efektif.

Berdasarkan hasil evaluasi *Confusion Matrix*, diperoleh nilai *True Positive (TP)* = 357 dan *True Negative (TN)* = 101 yang relatif tinggi, sehingga menunjukkan bahwa mayoritas data berhasil diklasifikasikan secara benar. Sementara itu, nilai *False Positive (FP)* = 29 dan *False Negative (FN)* = 7 tergolong rendah, yang mengindikasikan bahwa tingkat kesalahan klasifikasi masih dalam batas yang dapat diterima.

Distribusi sentimen yang tidak seimbang, yaitu dominasi kelas positif sebesar 80% dibandingkan kelas negatif sebesar 20%, turut mempengaruhi performa model. Model cenderung lebih optimal dalam mengklasifikasikan kelas positif dibandingkan kelas negatif. Hal ini terlihat dari nilai *precision*, *recall*, dan *F1-score* yang lebih tinggi pada kelas positif dibandingkan kelas negatif.

Berdasarkan hasil analisis terhadap ulasan pengguna aplikasi Grab, dominasi sentimen positif menunjukkan bahwa sebagian besar pengguna memberikan penilaian yang baik terhadap layanan yang diterima. Selain memberikan gambaran mengenai kecenderungan sentimen, hasil analisis juga mengindikasikan bahwa ulasan pengguna banyak membahas aspek-aspek layanan Grab, seperti kualitas layanan pengemudi, kemudahan penggunaan aplikasi, kecepatan pelayanan, serta pengalaman pengguna selama menggunakan aplikasi. Di sisi lain, ulasan yang memiliki sentimen negatif umumnya berkaitan dengan pengalaman pengguna yang kurang memuaskan terhadap layanan tertentu, sehingga dapat menjadi masukan bagi pihak Grab untuk melakukan evaluasi dan peningkatan kualitas layanan. Dengan demikian, hasil analisis sentimen tidak hanya menunjukkan proporsi ulasan positif dan negatif, tetapi juga memberikan informasi yang bermanfaat dalam mengidentifikasi aspek layanan yang menjadi perhatian utama pengguna.

Selain itu, keterbatasan pendekatan *Lexicon-Based* dalam pelabelan awal juga menjadi salah satu faktor yang mempengaruhi hasil klasifikasi. Kamus sentimen yang digunakan belum sepenuhnya mampu menangkap variasi bahasa informal, seperti penggunaan slang, singkatan, maupun konteks negasi dalam ulasan pengguna. Akibatnya, beberapa ulasan negatif masih berpotensi terklasifikasi sebagai positif.

Meskipun demikian, kombinasi metode *Lexicon-Based* sebagai pelabelan otomatis dan *KNN* sebagai classifier terbukti efektif untuk analisis sentimen teks berbahasa Indonesia. Pendekatan ini memberikan solusi yang sederhana, efisien, dan tidak memerlukan pelabelan manual dalam jumlah besar, sehingga untuk penelitian dengan keterbatasan dataset.

KESIMPULAN

Penelitian ini berhasil menerapkan kombinasi metode *Lexicon-Based* dan algoritma *KNN* untuk melakukan analisis sentimen terhadap ulasan pengguna aplikasi Grab pada platform Google Play Store. Dataset yang digunakan dalam penelitian ini berjumlah 2.467 ulasan berbahasa Indonesia yang melalui beberapa tahapan pengolahan data, meliputi preprocessing, tokenizing, stopword removal, stemming, pelabelan sentimen menggunakan pendekatan *Lexicon-Based*, ekstraksi fitur menggunakan *TF-IDF*, serta proses klasifikasi menggunakan algoritma *KNN*. Berdasarkan hasil analisis sentimen yang dilakukan, distribusi data didominasi oleh sentimen positif sebesar 80%, sedangkan sentimen negatif sebesar 20%, yang menunjukkan bahwa sebagian besar pengguna memberikan tanggapan positif terhadap layanan aplikasi Grab. Hasil pengujian model klasifikasi menggunakan algoritma *KNN* memperoleh nilai *accuracy* sebesar 92,71%, *precision* sebesar 92,48%, *recall* sebesar 98,08%, dan *F1-score* sebesar 95,20%, sehingga menunjukkan bahwa metode yang digunakan mampu mengklasifikasikan sentimen ulasan pengguna dengan performa yang baik. Selain itu, penerapan metode *Lexicon-Based* pada proses pelabelan otomatis dinilai dapat membantu mengurangi proses pelabelan manual yang membutuhkan waktu lebih lama. Dengan demikian, kombinasi metode *Lexicon-Based* dan *KNN* terbukti efektif untuk diterapkan dalam analisis sentimen ulasan aplikasi transportasi online berbahasa Indonesia serta dapat digunakan sebagai bahan evaluasi layanan berdasarkan opini pengguna yang diberikan melalui Google Play Store.

Penelitian ini tidak hanya memberikan kontribusi praktis dalam penerapan metode *Lexicon-Based* dan *KNN* untuk analisis sentimen ulasan aplikasi Grab berbahasa Indonesia, tetapi juga memberikan kontribusi ilmiah terhadap pengembangan penelitian analisis sentimen melalui implementasi pelabelan sentimen otomatis menggunakan kamus sentimen bahasa Indonesia (*InSet*) yang dipadukan dengan representasi fitur *TF-IDF* dan algoritma *KNN*. Pendekatan tersebut menunjukkan bahwa pelabelan sentimen otomatis dapat menjadi alternatif yang efektif untuk mengurangi ketergantungan terhadap proses anotasi manual pada penelitian analisis sentimen berbahasa Indonesia. Selain itu, hasil penelitian ini memiliki implikasi praktis bagi pengembang aplikasi maupun pengambil keputusan dalam memanfaatkan informasi sentimen pengguna sebagai dasar evaluasi kualitas layanan, penyusunan strategi peningkatan kepuasan pengguna, serta pengambilan keputusan yang lebih berbasis data.

DAFTAR PUSTAKA

- Amal, I., & Jayanta. (2023). Perbandingan Pelabelan Otomatis Dan Manual Untuk Analisis Sentimen Terhadap Kenaikan Harga BBM Pertamina Pada Twitter Menggunakan Algoritma Support Vector Machine. *Seminar Nasional Mahasiswa Ilmu Komputer Dan Aplikasinya (SENAMIKA)*, 4(2), 473–487.
- Angelina, S. J., Bijaksana, A., Negara, P., & Muhandi, H. (2023). Analisis Pengaruh Penerapan Stopword Removal Pada Performa Klasifikasi Sentimen Tweet Bahasa Indonesia. *JUARA (Jurnal Aplikasi Dan Riset Informatika)*, 02(1), 165–173.
<https://doi.org/10.26418/juara.v2i1.69680>
- Dyo, F. A. H., Hayatin, N. H., & Aditya, C. S. K. (2020). Analisa Sentimen Tweet Berbahasa Indonesia Dengan Menggunakan Metode Lexicon Pada Topik Perpindahan Ibu Kota Indonesia. *Jurnal Repositor*, 2(7), 977.
<https://doi.org/10.22219/repositor.v2i7.937>
- Legito, L., Riau, N. P., Putro, A. N. S., Mardiani, E., Arifin, N. Y., Sepriano, S., & Erkamim, M. (2023). Penerapan Algoritma K-Nearest Neighbor untuk Analisis Sentimen Terhadap Isu Khilafah dan Radikalisme di Indonesia: Implementation K-Nearest Neighbor Algorithm for Sentiment Analysis on Khilafah and Radicalism Issues in Indonesia. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 3(2), 324–330.
- Maheri, R., Salisah, F. N., Muttakin, F., & Megawati, M. (2025). Analisis Sentimen Ulasan Aplikasi M-Paspor Menggunakan Naive Bayes Dan Support Vector Machine. *JUPI (Jurnal Ilmiah Penelitian Dan Pembelajaran Informatika)*, 10(1), 448–458.
<https://doi.org/10.29100/jupi.v10i1.5826>
- Maldini, M. A. A., & Septi, A. (2025). Analisis Sentimen Ulasan Pengguna Aplikasi Flip. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 9(3), 4098–4105.
- Pardede, J. (2025). Perbandingan Algoritma Stemming Porter, Sastrawi, Idris, Comparison of Stemming Algorithms Porter, Sastrawi, Idris, and Arifin Setiono on Indonesian Text Documents. *Jurnal Teknologi Informasi Dan Ilmu Komputer (JTIK)*, 12(1).
<https://doi.org/10.25126/jtiik.2025128860>
- Purnamasari, D., Aji, A. B., Wulandari, D., Reza, F. A., Safirla, M., Yanda, N., & Hidayati, U. (2023). Pengantar Metode Analisis Sentimen. In *Pengantar Metode Analisis Sentimen*.
- Riansyah, G. R., & Sugianto, C. A. (2024). Analisis Sentimen Ulasan Aplikasi Dana Di Google Play Store Menggunakan Algoritma Naive Bayes. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(5), 9849–9857.
<https://doi.org/10.36040/jati.v8i5.10732>
- Ridwan, M. U., & Rais, A. R. (2025). Analisis Sentimen Ulasan Pengguna Terhadap Kualitas Layanan Transportasi Online Pada Aplikasi Kai Access Menggunakan Metode Deep Learning Dengan Algoritma Lstm, Gru, Dan Cnn. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 9(5), 7546–7555.
<https://doi.org/10.36040/jati.v9i5.14744>
- Rifaldi, D., & Fadlil, A. (2023). Teknik preprocessing pada text mining menggunakan data tweet “mental health”. *DECODE: Jurnal Pendidikan Teknologi Informasi*, 3(2), 161–171.
- Setyani, T., Sari, K., Heidy, H. N., & Suryono, R. R. (2026). Analisis Sentimen Pengguna Aplikasi Jamsostek Mobile Berdasarkan Ulasan Google Play Store Menggunakan Algoritma Support. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 6(1), 373–384.
- Sukestiyarno, Y. L., & Agoestanto, A. (2017). Batasan prasyarat uji normalitas dan uji homogenitas pada model regresi linear. *Unnes Journal of Mathematics*, 6(2), 168–177.
- Syahputri, R., Tanubrata, J. K., & Trisnawati, S. (2024). Penerapan Lexicon-Based Untuk Analisis Sentiment Masyarakat Terhadap Kasus Kebocoran Data di Indonesia. *Journal AI2MTEch*, 1(2), 93–105.
- Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). Lexicon-based methods for sentiment analysis. *Computational Linguistics*, 37(2), 267–307.
https://doi.org/10.1162/COLI_a_00049
- Tribuana, D., Usman, U., & Dayanti, D. (2025). Penerapan Natural Language Processing Untuk Analisis Sentimen Terhadap Layanan Publik Di Media Sosial Twitter. *Jurnal Teknologi Dan Bisnis Cerdas*, 1(1), 28–37.
<https://doi.org/10.64476/jtbc.v1i1.3>