

**ANALISIS SENTIMEN PENGGUNA KERETA REL LISTRIK
(KRL) JOGJA SOLO BERDASARKAN ULASAN SOSIAL
MEDIA MENGGUNAKAN METODE NAÏVE BAYES**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

MUH. FAJRI KANUGRAHA AGUNG

21.11.4137

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

**ANALISIS SENTIMEN PENGGUNA KERETA REL LISTRIK
(KRL) JOGJA SOLO BERDASARKAN ULASAN SOSIAL
MEDIA MENGGUNAKAN METODE NAÏVE BAYES**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

MUH. FAJRI KANUGRAHA AGUNG

21.11.4137

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

HALAMAN PERSETUJUAN

SKRIPSI

ANALISIS SENTIMEN PENGGUNA KERETA REL LISTRIK (KRL)

JOGJA SOLO BERDASARKAN ULASAN SOSIAL MEDIA

MENGGUNAKAN NAÏVE BAYES

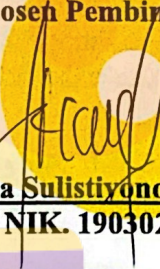
yang disusun dan diajukan oleh

Muh. Fajri Kanugraha Agung

21.11.4137

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 25 Agustus 2025

Dosen Pembimbing,


Mulia Sulistiyono, M.Kom

NIK. 190302057

HALAMAN PENGESAHAN

SKRIPSI

**ANALISIS SENTIMEN PENGGUNA KERETA REL LISTRIK (KRL)
JOGJA-SOLO BERDASARKAN ULASAN MEDIA SOSIAL
MENGUNAKAN NAÏVE BAYES**

yang disusun dan diajukan oleh

Muh. Fajri Kanugraha Agung

21.11.4137

Telah dipertahankan di depan Dewan Penguji
pada tanggal 25 Agustus 2025

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Hanif Al Fatta, S.Kom., M.Kom., Ph.D.
NIK. 190302096

Ike Verawati, S.Kom., M.Kom.
NIK. 190302237

Mulia Sulistiyono, S.Kom., M.Kom.
NIK. 190302248

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal < tanggal lulus ujian >

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusriani, M.Kom.
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Muh. Fajri Kanugraha Agung
NIM : 21.11.4137

Menyatakan bahwa Skripsi dengan judul berikut:

**Analisis Sentimen Pengguna Kereta Rel Listrik (KRL) Jogja Solo
berdasarkan Media Sosial Menggunakan Naïve Bayes**

Dosen Pembimbing : Mulia Sulistiyono, M.Kom,

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 25 Agustus 2025

Yang Menyatakan,



Muh. Fajri Kanugraha Agung

HALAMAN PERSEMBAHAN

Terwujudnya Skripsi ini tidak akan selesai tanpa bantuan dari berbagai pihak yang terus mendorong peneliti untuk menyelesaikannya, Oleh karena itu, peneliti mengucapkan terimakasih kepada :

1. Mulia Sulistiyono, S.Kom, M.Kom selaku Dosen Pembimbing yang telah memberikan bimbingan, arahan, serta masukan yang sangat berharga selama proses penyusunan skripsi ini. Kesabaran dan dukungan beliau menjadi motivasi besar bagi peneliti dalam menyelesaikan penelitian ini.
2. Keluarga peneliti, khususnya Abah dan Ibu tercinta, yang selalu memberikan doa, kasih sayang, serta dukungan moril dan materil tanpa henti. Doa dan pengorbanan yang diberikan menjadi kekuatan utama bagi peneliti dalam menjalani setiap proses.
3. Teman-teman DarkSide, yang selalu hadir memberikan semangat, kebersamaan, diskusi, serta dukungan baik dalam suka maupun duka selama masa studi dan penyusunan skripsi. Kehadiran kalian sangat berarti bagi peneliti.
4. Terakhir, peneliti menyampaikan terima kasih kepada diri sendiri, yang berhasil melalui semua ini. Terimakasih sudah bertahan hingga akhir dan selalu menyelesaikan apa yang sudah dimulai.

KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Allah SWT atas segala limpahan rahmat, taufik, dan hidayah-Nya sehingga penulis dapat menyelesaikan penyusunan skripsi ini dengan baik. Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana pada Program Studi Informatika, Universitas Amikom Yogyakarta.

Selama proses penyusunan skripsi ini, penulis mendapatkan banyak dukungan, arahan, dan bimbingan dari berbagai pihak. Ucapan terima kasih penulis sampaikan kepada Bapak Mulia Sulistiyono, S.Kom., M.Kom. selaku Dosen Pembimbing yang dengan sabar memberikan bimbingan, arahan, serta masukan yang sangat berharga selama penelitian berlangsung. Penulis juga menghaturkan terima kasih yang tak terhingga kepada Abah dan Ibu tercinta yang selalu memberikan doa, kasih sayang, serta dukungan moril dan materil yang tiada henti, sehingga menjadi sumber kekuatan dalam menyelesaikan studi ini. Tidak lupa, penulis juga menyampaikan rasa terima kasih kepada teman-teman seperjuangan yang selalu hadir memberikan semangat, motivasi, serta kebersamaan selama masa studi hingga penyusunan skripsi ini.

Penulis menyadari sepenuhnya bahwa skripsi ini masih jauh dari sempurna. Oleh karena itu, kritik dan saran yang membangun sangat penulis harapkan demi perbaikan di masa mendatang. Semoga karya ini dapat memberikan manfaat bagi pembaca, memperluas wawasan dalam bidang analisis sentimen, serta menjadi referensi bagi penelitian selanjutnya.

Yogyakarta, 25 Agustus 2025

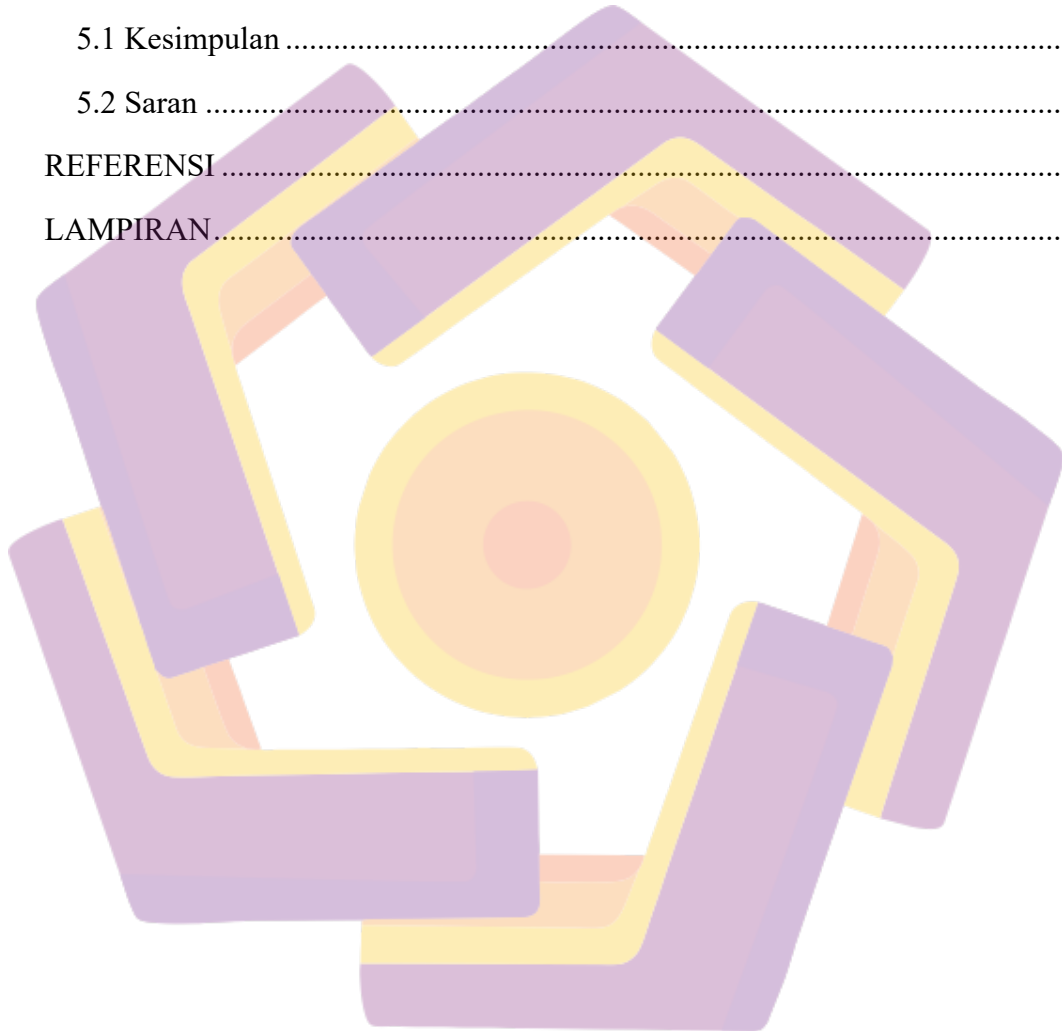
Penulis

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR TABEL.....	x
DAFTAR GAMBAR	xi
DAFTAR LAMPIRAN.....	xii
DAFTAR LAMBANG DAN SINGKATAN	xiii
DAFTAR ISTILAH	xiv
INTISARI	xv
<i>ABSTRACT</i>	xvi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	2
1.3 Batasan Masalah	2
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.6 Sistematika Penulisan	4
BAB II TINJAUAN PUSTAKA	5
2.1 Studi Literatur	5
2.2 Dasar Teori.....	10
2.2.1 Media Sosial.....	10
2.2.2 Data Mining	10
2.2.3 Python	10

2.2.4 Analisis Sentimen	11
2.2.5 Naïve Bayes	11
2.2.6 Term Frequency – Inverse Document Frequency (TF-IDF).....	12
2.2.7 Confussion Matrix	13
BAB III METODE PENELITIAN	16
3.1 Alur Penelitian	16
3.1.1 Pengumpulan Data	17
3.1.2 Praproses Data (Data Preprocessing).....	18
3.1.3 Modeling.....	19
3.1.4 Evaluasi.....	22
3.2 Alat dan Bahan.....	22
BAB IV HASIL DAN PEMBAHASAN	24
4.1 Pengumpulan Data	24
4.1.1 Scraping Data.....	24
4.1.2 Penggabungan Data	25
4.2 Praprosesing Data (<i>Preprocessing Data</i>).....	26
4.2.1 Case Folding	26
4.2.2 Text Cleaning.....	27
4.2.3 Tokenizing	27
4.2.4 Normalization Slang Words.....	28
4.2.5 Stopwords Removal	29
4.2.6 Stemming	30
4.2.7 Noise Word Removal.....	31
4.2.8 Rejoin Token.....	31
4.3 Modeling.....	32
4.3.1 Labeling	32
4.3.2 Data Splitting	34

4.3.3 Ekstraksi Fitur TF-IDF	34
4.3.4 SMOTE <i>Oversampling</i>	36
4.3.5 Naïve Bayes	37
4.4 Evaluasi Niave Bayes Berdasarkan Confusion Matrix	38
BAB V PENUTUP	40
5.1 Kesimpulan	40
5.2 Saran	41
REFERENSI	42
LAMPIRAN	44

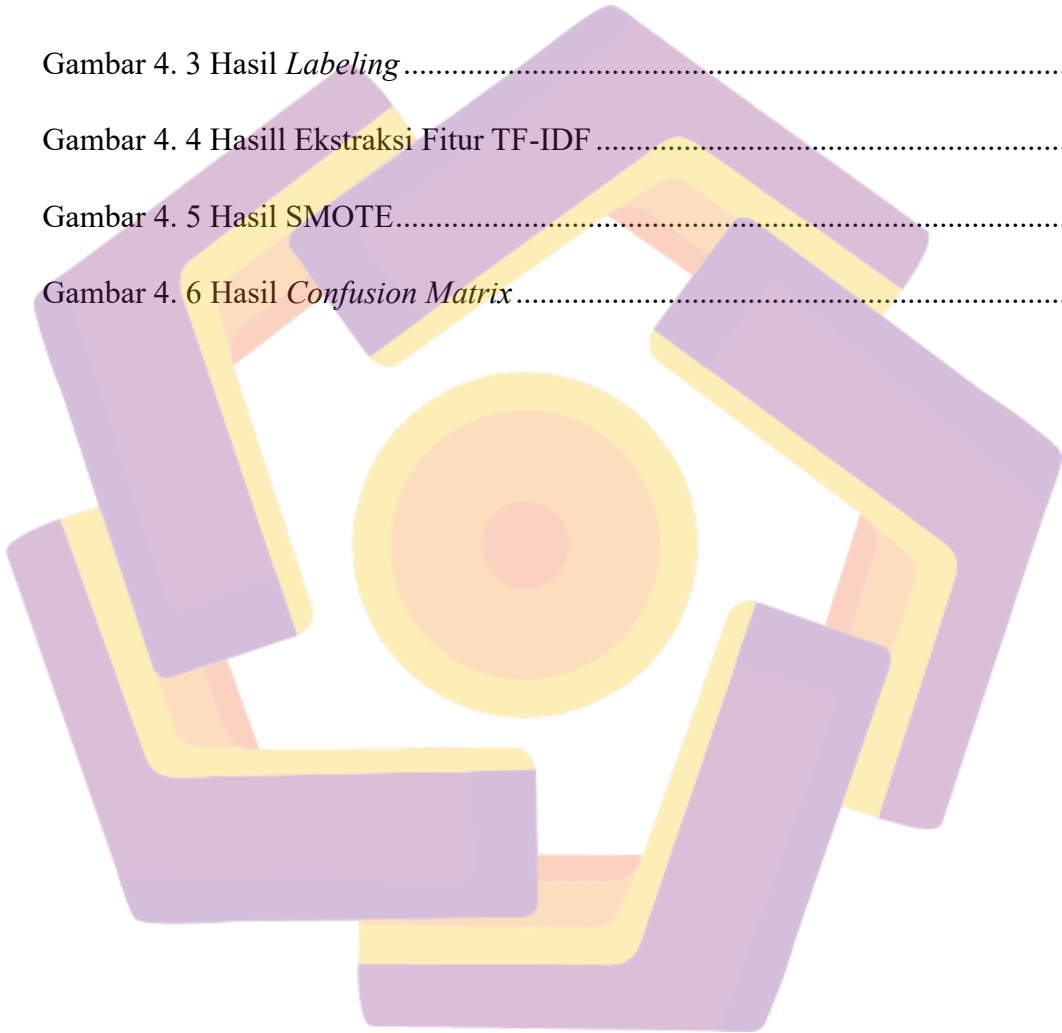


DAFTAR TABEL

Tabel 2. 1 Keaslian Penelitian.....	7
Tabel 2. 2 <i>Consufion Matrix</i>	13
Tabel 4. 1 Hasil <i>Case Folding</i>	26
Tabel 4. 2 Hasil <i>Text Cleaning</i>	27
Tabel 4. 3 Hasil <i>Tokenizing</i>	28
Tabel 4. 4 Hasil <i>Normalization Slang Words</i>	28
Tabel 4. 5 Hasil <i>Stopword Removal</i>	29
Tabel 4. 6 Hasil <i>Stemming</i>	30
Tabel 4. 7 Hasil <i>Noise Word Removal</i>	31
Tabel 4. 8 Hasil <i>Rejoin Token</i>	32
Tabel 4. 9 Frekuensi Kata Setelah Labeling	33
Tabel 4. 10 Hasil Evaluasi Model pada Data Uji.....	37

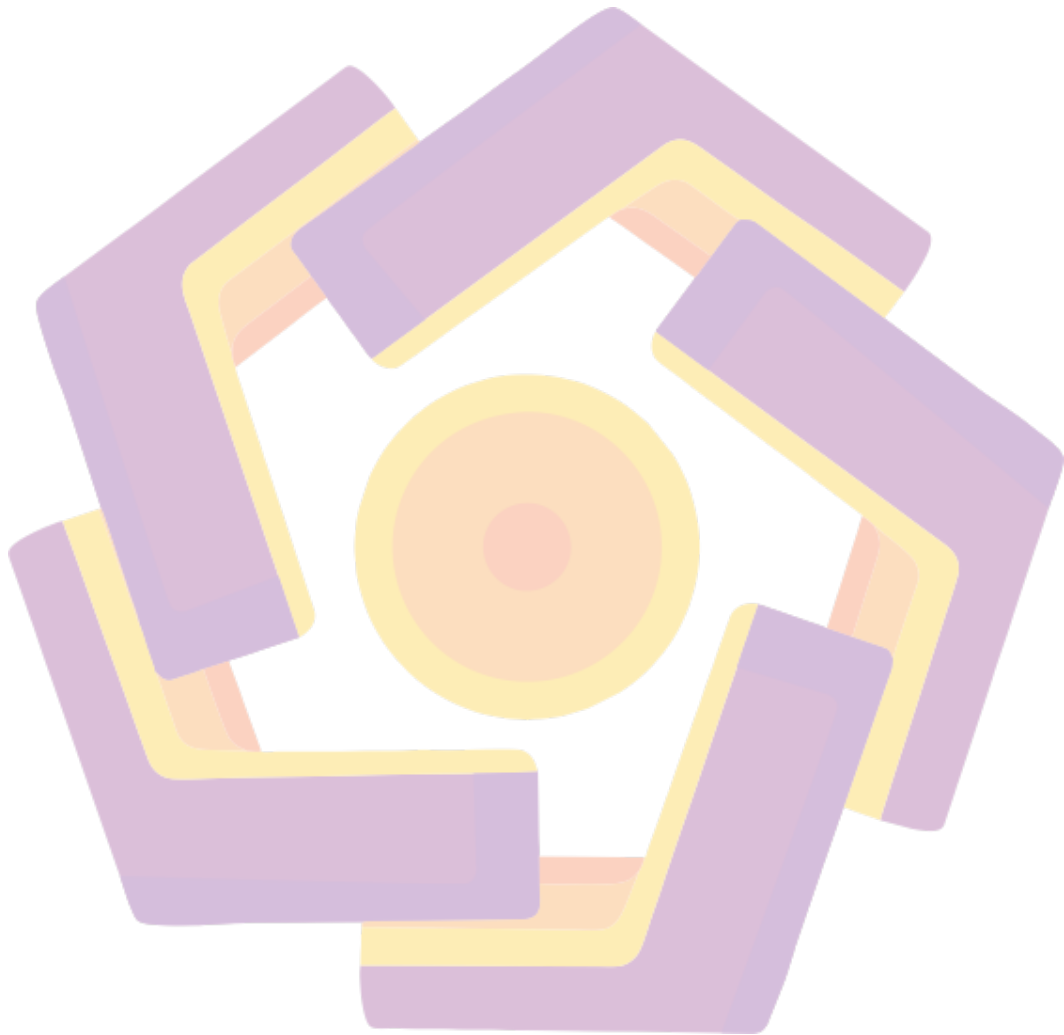
DAFTAR GAMBAR

Gambar 3. 1 Alur Penelitian	16
Gambar 4. 1 Data ulasan <i>Twitter</i>	25
Gambar 4. 2 Data ulasan <i>TikTok</i>	25
Gambar 4. 3 Hasil <i>Labeling</i>	34
Gambar 4. 4 Hasil Ekstraksi Fitur TF-IDF	35
Gambar 4. 5 Hasil SMOTE	36
Gambar 4. 6 Hasil <i>Confusion Matrix</i>	38



DAFTAR LAMPIRAN

Lampiran 1. Profil obyek Penelitian	10
Lampiran 2. Dokumentasi Penelitian	11



DAFTAR LAMBANG DAN SINGKATAN

KRL	Kereta Rel Listrik
TF-IDF	Term Frequency – Inverse Document Frequency
NB	Naïve Bayes
KNN	K-Nearest Neighbor
NLP	Natural Language Processing
SMOTE	Synthetic Minority Oversampling Technique
TP	True Positive
TN	True Negative
FP	False Positive
FN	False Negative
Σ	Sigma
log	Logaritma

DAFTAR ISTILAH

Vektor	besaran yang mempunyai arah
Eigen Value	akar akar persamaan
Tuning	Mengoptimalkan hyperparameter
Klasifikasi	Mengelompokkan data ke kelas tertentu
Parameter	Nilai dalam model yang diatur saat pelatihan
Sentimen	Analisis emosi dalam teks
Naïve Bayes	Algoritma klasifikasi probabilistic
Supervised Learning	Pembelajaran dengan data berlabel
Hybrid	Gabungan Metode
Vektor	Representasi nilai dengan arah
Support Vectors	Titik data penentu margin
Library	Kumpulan kode siap pakai
Word Embedding	Representasi kata sebagai vector
Word2vec	Algoritma untuk embedding kata
Model	Representasi matematis untuk analisis
Input	Data masukan ke model
Error Analysis	Analisis kesalahan prediksi model
Data Mining	Menemukan pola dalam data
Confusion Matrix	Tabel evaluasi model
Oversampling	Teknik penyeimbangan data minoritas
Preprocessing	Pemrosesan awal data
Lowercase	Mengubah teks jadi huruf kecil
Slang	Bahasa informal
Machine Learning	Mesin belajar dari data
Laplace Smoothing	Teknik smoothing untuk menghindari probabilitas nol
Alpha	Parameter smoothing pada Naïve Bayes
Lexicon	Kumpulan kosakata/kamus beserta informasi tambahan

INTISARI

KRL Commuter Line Jogja–Solo yang dioperasikan sejak 10 Februari 2021 sebagai pengganti Kereta Api Prambanan Ekspres (Prameks) hadir untuk meningkatkan layanan transportasi umum di wilayah Yogyakarta–Solo. Meskipun menawarkan tarif terjangkau dan perjalanan efisien, layanan ini masih mendapat keluhan pengguna, terutama pada jam sibuk yang menimbulkan kepadatan penumpang, sehingga memicu opini beragam di media sosial. *Twitter* dan *TikTok* menjadi sumber utama ulasan masyarakat secara real-time, yang dapat dimanfaatkan untuk menganalisis persepsi publik. Penelitian ini menggunakan algoritma *Multinomial Naïve Bayes* untuk mengklasifikasikan sentimen pengguna KRL Jogja–Solo ke dalam tiga kategori: positif, netral, dan negatif. Data berupa teks opini dikumpulkan dari *Twitter* dan *TikTok*, melalui tahapan praproses teks, ekstraksi fitur dengan TF-IDF, penyeimbangan kelas menggunakan SMOTE, pemodelan dengan *Naïve Bayes*, dan evaluasi menggunakan *confusion matrix*. Hasil penelitian menunjukkan bahwa sentimen netral mendominasi ulasan, dengan akurasi model sebesar 76,21% dan *f1-score* tertinggi pada kelas netral sebesar 0,85, sedangkan positif dan negatif masing-masing 0,49 dan 0,45. Meskipun distribusi kinerja antar kelas belum merata, hasil ini membuktikan bahwa *Naïve Bayes* masih relevan digunakan untuk analisis sentimen sederhana dan efisien. Penelitian ini dapat dimanfaatkan oleh penyedia layanan transportasi untuk mengevaluasi kualitas layanan berdasarkan data opini publik di media sosial.

Kata kunci: Analisis Sentimen, *Naïve Bayes*, KRL Jogja–Solo, *Twitter*, *TikTok*.

ABSTRACT

The Yogyakarta–Solo Commuter Line (KRL Jogja–Solo), which began operation on February 10, 2021, replacing the Prambanan Ekspres (Prameks), was introduced to improve public transportation services in the Yogyakarta–Solo area. Although it offers affordable fares and efficient travel, the service has received user complaints, particularly during peak hours when passenger congestion occurs, leading to diverse opinions on social media. Twitter and TikTok serve as the main platforms for real-time public reviews, providing valuable data to analyze public perception. This study employs the Multinomial Naïve Bayes algorithm to classify user sentiments toward the KRL Jogja–Solo into three categories: positive, neutral, and negative. Textual opinion data were collected from Twitter and TikTok, followed by text preprocessing, feature extraction using TF-IDF, class balancing with SMOTE, sentiment classification using Naïve Bayes, and evaluation via a confusion matrix. The results show that neutral sentiment dominates user reviews, with the model achieving an accuracy of 76.21% and the highest f1-score for the neutral class at 0.85, while positive and negative sentiments achieved 0.49 and 0.45, respectively. Although class performance distribution remains uneven, the findings demonstrate that Naïve Bayes remains relevant for simple and efficient sentiment analysis tasks. This research can be utilized by transportation service providers to evaluate service quality based on public opinion data from social media.

Keyword: Sentiment Analysis, Naïve Bayes, KRL Jogja–Solo, Twitter, TikTok.