

ANALISIS SENTIMEN SISTEM E-TILANG PADA PLATFORM TWITTER MENGGUNAKAN METODE NAIVE BAYES

Shella Ayudya Aryani Armand, Moch Hafid T, Muhammad Rafi Muttaqin

Program Studi Teknik Informatika S1, STT Wastukencana Purwakarta
Jalan Cikopak No. 53, Mulyamekar, Kec.Babakancikao, Kab. Purwakarta, Jawa Barat
shellaayudya06@wastukencana.ac.id

ABSTRAK

Berita yang belakangan ini menjadi sorotan di media sosial Twitter di Indonesia adalah penerapan sistem e-Tilang, yang merupakan langkah baru dari pemerintah dalam mewujudkan Good Governance. Sistem ini diharapkan dapat mendisiplinkan para pengendara kendaraan bermotor yang sering melakukan pelanggaran berlalu lintas. Banyak masyarakat memiliki pendapat yang beragam mengenai penerapan sistem ini, ada yang setuju dan ada yang tidak setuju. Dalam penelitian ini, dilakukan analisis sentimen terhadap sistem e-Tilang atau pendapat masyarakat untuk mengelompokkan komentar-komentar tersebut menjadi kesan positif, netral, dan negatif. Algoritma Naive Bayes dipilih untuk analisis karena memiliki nilai probabilitas atau peluang tertinggi untuk pengklasifikasian data. Proses pelabelan data pada Twitter menggunakan Bahasa Pemrograman Python dan Lexicon Based dengan menggunakan tools Google Colab. Data kemudian melewati empat tahapan preprocessing, yaitu transformation, tokenizing, filtering, dan stemming. Setelah data selesai dipreprocessing, dilakukan pembobotan TF-IDF untuk mengetahui banyaknya kemunculan kata dalam setiap komentar. Hasil dari pengklasifikasian data menggunakan Algoritma Naive Bayes dievaluasi menggunakan Confusion Matrix pada tools Google Colab. Hasil evaluasi menunjukkan bahwa presisi positif sebesar 43%, presisi negatif sebesar 38%, dan presisi netral sebesar 85%. Sementara itu, recall positif sebesar 43%, recall negatif 14%, dan recall netral sebesar 95%, dengan akurasi keseluruhan mencapai 80%. Dari hasil analisis ini, dapat disimpulkan bahwa tanggapan masyarakat mengenai sistem e-Tilang yang diterapkan oleh pemerintah Indonesia cenderung netral.

Kata kunci: analisis sentimen, e-Tilang, Naive Bayes, Twitter, Python

1. PENDAHULUAN

Dalam kehidupan sosial, mengungkapkan sentimen kepada orang lain menjadi kegiatan sehari-hari. Kemajuan teknologi juga memungkinkan masyarakat mengungkapkan pendapat melalui media sosial seperti Twitter. Twitter merupakan jejaring sosial terbesar di dunia dan telah menghasilkan keuntungan mencapai USD 145 juta [1].

Saat ini, di Twitter khususnya di Indonesia, permasalahan yang sedang menjadi perbincangan adalah penerapan sistem e-Tilang [1]. Sistem e-Tilang merupakan langkah baru yang diambil pemerintah Indonesia dalam mewujudkan Good Governance, tetapi penerapannya menimbulkan pro dan kontra. Beberapa pihak mendukung sistem ini karena menganggapnya dapat mencegah kenakalan penegak hukum dari pungutan liar, menciptakan ketertiban, dan meningkatkan akuntabilitas uang denda. Namun, sebagian pihak tidak setuju karena merasa sistem e-Tilang masih belum optimal, seperti mudahnya pemalsuan atau pencopotan nomor polisi, dan adanya beberapa laporan tentang kesalahan tilang yang terjadi. [2].

Analisis sentimen adalah studi yang bertujuan untuk menganalisis opini, sentimen, dan emosi yang terdapat dalam dokumen atau data. Text mining merupakan proses otomatis untuk teks yang bertujuan untuk mendapatkan informasi yang berguna dari sekumpulan dokumen atau mencari kata-kata yang dapat mewakili isi dokumen sehingga memungkinkan

analisis keterhubungan antar dokumen[3]. Penulis memilih metode Naïve Bayes Classifier dalam text mining karena metode ini sering bekerja lebih baik dalam banyak situasi dunia nyata yang kompleks. Naïve Bayes Classifier merupakan salah satu algoritma klasifikasi yang berbasis probabilitas dan dapat digunakan untuk melakukan analisis sentimen dengan memprediksi sentimen positif, negatif, atau netral dari teks.

Penelitian ini sangat membantu pemerintah dalam melakukan evaluasi terhadap program yang dijalankan, khususnya terkait dengan sistem e-Tilang. Dengan melakukan analisis sentimen masyarakat terhadap sistem e-Tilang, pemerintah dapat mendapatkan informasi yang berharga tentang pandangan dan reaksi masyarakat terhadap implementasi sistem ini. Tanpa adanya evaluasi dari masyarakat, pemerintah akan kesulitan mengetahui sejauh mana keberhasilan sistem e-Tilang yang dijalankan serta memahami aspek-aspek yang perlu diperbaiki untuk meningkatkan efektivitasnya. Dengan demikian, penelitian ini dapat menjadi landasan penting dalam pengambilan keputusan dan peningkatan program yang sedang berjalan. [4].

2. TINJAUAN PUSTAKA

2.1. Analisis Sentimen

Analisis sentimen menentukan sentimen dalam teks, apakah positif, negatif, atau netral. Populer dan berdampak dalam berbagai bidang kehidupan [5].

2.2. Sistem E-Tilang

E-Tilang adalah tilang elektronik untuk meningkatkan efektivitas dan efisiensi penegakan hukum lalu lintas. Masyarakat bisa membayar denda melalui aplikasi setelah data tercatat. Ini membuat proses tilang lebih transparan dan terencana [6].

2.3. Twitter

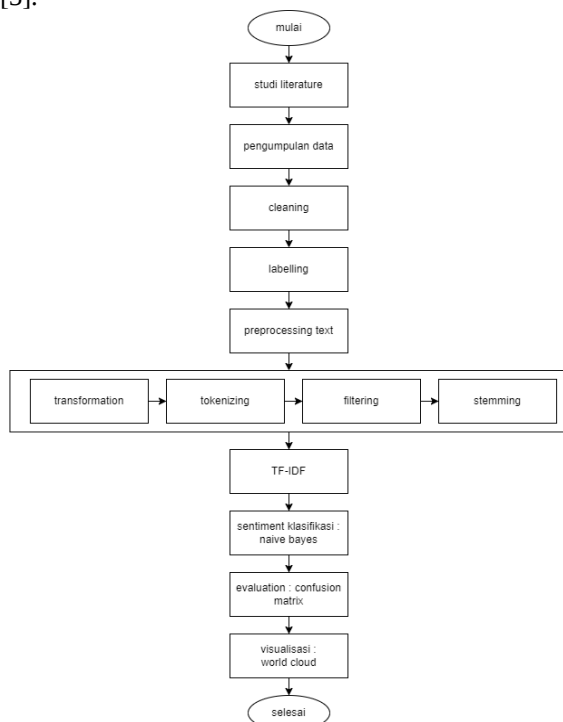
Twitter adalah platform media sosial di mana orang dari seluruh dunia dapat mengungkapkan pendapat mereka. Data yang dihasilkan di *Twitter* memiliki potensi besar untuk dianalisis, karena dapat diekstrak menjadi informasi penting melalui analisis opini (opinion mining). Opini mengenai berita, peluncuran produk, atau tren tertentu dapat diamati dengan baik di *Twitter* [7].

2.4. Naive Bayes

Naive Bayes adalah algoritma pengklasifikasi probabilistik sederhana yang menghitung probabilitas berdasarkan frekuensi dan kombinasi nilai dataset. Algoritma ini mengasumsikan atribut independen jika nilai variabel kelas diketahui [8].

2.5. Text Mining

Text Mining adalah metode yang mencari informasi penting dari dokumen teks melalui teknik preprocessing, pencarian, ekstraksi data, dan pengkategorian. Ini adalah cabang analisis data dan pemrosesan bahasa alami, yang tujuannya adalah mengenali pola dan keterkaitan kata-kata dalam teks untuk mengekstrak informasi bermakna. Metode dalam Text Mining termasuk klasifikasi, analisis sentimen, ekstraksi entitas, dan penambangan topik [3].



Gambar 1. Kerangka Pemikiran

3. METODE PENELITIAN

Penelitian menggunakan data dari sistem e-Tilang di Indonesia dalam periode 09 Desember 2021 hingga 24 Mei 2023. Setelah cleaning, data berkurang menjadi 1.696 tweet dengan label 3 kelas: positif (150 data), netral (1.453 data), dan negatif (93 data).

3.1. Kerangka Pemikiran

Proses ini mengubah data mentah menjadi informasi yang dapat diklasifikasikan ke dalam kelas tertentu setelah pengolahan data. Berikut merupakan kerangka pemikiran yang dapat dilihat pada Gambar 1.

3.2. Studi Literature

Tahap awal penelitian adalah melakukan review jurnal-jurnal sebelumnya yang menggunakan metode Naive Bayes untuk analisis sentimen. Hasil review menjadi referensi dan pedoman untuk merumuskan tahapan penelitian yang akan dilakukan, termasuk memilih dan mengadaptasi pendekatan dan metode dari penelitian terdahulu sesuai dengan tujuan penelitian yang ingin dicapai.

3.3. Crawling Data

Penelitian ini menggunakan dataset dari Twitter yang dikumpulkan melalui proses pencarian data menggunakan NodeJS dan Command Prompt. Data berupa tweet dari platform Twitter. NodeJS adalah platform untuk menjalankan kode JavaScript di sisi server atau backend.

3.4. Cleaning

Tahap cleaning membersihkan tweet dari kata-kata tidak dibutuhkan, menghapus data duplikat dan kosong. Jumlah data berkurang dari 1.882 tweet menjadi 1.696 tweet. Berikut merupakan contoh hasil dari tahapan *cleaning* dapat dilihat pada Gambar 2.

Wapada Surat E-Tilang dengan format APK (aplikasi). Pengiriman surat E-tilang hanya diberikan melalui PT Pos Indonesia ke alamat penagang yang sesuai karena di sini jika menerima surat e-tilang selain dari pengiriman pos, seperti Via WhatsApp harap abaikan kemp, pengembalian.	Wapada Surat E-Tilang dengan format APK (aplikasi). Pengiriman surat E-tilang hanya diberikan melalui PT Pos Indonesia ke alamat penagang yang sesuai karena di sini jika menerima surat e-tilang selain dari pengiriman pos, seperti Via WhatsApp harap abaikan kemp, pengembalian.
#Ditantas #PoldaAceh #E-Tilang Ditantas Berakut E-Tilang	Ditantas PoldaAceh E-Tilang Ditantas Berakut E-Tilang
#Etilang #BandaAceh #SuratTilang Direktorat Lalu Lintas Ditantas Polda Aceh resmi menerapkan sistem e-tilang lalu lintas di Kota Banda Aceh	Etilang BandaAceh SuratTilang Direktorat Lalu Lintas Ditantas Polda Aceh resmi menerapkan sistem e-tilang lalu lintas di Kota Banda Aceh
#Foto Jalan Riusna Sidi, Jakarta, kembali rusak cukup parah. Banyak lubang di sepanjang jalan. Foto: Agung Pambudi@yellcom	Foto Jalan Riusna Sidi, Jakarta, kembali rusak cukup parah. Banyak lubang di sepanjang jalan. Foto: Agung Pambudi@yellcom
#KurovRek [02/08/2022] Sosialisasi Sistem E-tilang di Kecamatan Punorejo. Senin (01/08/2022), pegawai lingkungan Kecamatan Punorejo mendapatkan sosialisasi terkait sistem kerja e-tilang serta bagaimana proses pembayaran e-tilang.	KurovRek Sosialisasi Sistem Etilang di Kecamatan Punorejo. Senin pegawai lingkungan Kecamatan Punorejo mendapatkan sosialisasi terkait sistem kerja e-tilang serta bagaimana proses pembayaran e-tilang.
#MostPopular Wapada Ada Modus Penipuan Berkedok Aplikasi E-Tilang	MostPopular Wapada Ada Modus Penipuan Berkedok Aplikasi E-Tilang
#Nawakgalem. Polresta Malang Kota berencana menerapkan e-tilang pada bulan Mei 2021 untuk menekan dan mencegah angka kecelakaan serta meningkatkan kedisiplinan berlalu lintas. Ada 4 titik yg akan dipasang kamera sebagaimana disebut dalam infografs.	Nawakgalem Polresta Malang Kota berencana menerapkan e-tilang pada bulan Mei 2021 untuk menekan dan mencegah angka kecelakaan serta meningkatkan kedisiplinan berlalu lintas. Ada 4 titik yg akan dipasang kamera sebagaimana disebut dalam infografs.
#Report @divsihumaspti Poin Libatkan PT Pos Untuk Optimalkan E-Tilang (ETLE) #PoliPress	Report Poin Libatkan PT Pos Untuk Optimalkan E-Tilang (ETLE) PoinPress

Gambar 2. Hasil *Cleaning*

3.5. Labelling

Proses labelling menetapkan sentimen pada setiap tweet menggunakan library lexicon based. Library ini berisi daftar kata-kata dengan nilai sentimen untuk klasifikasi ke dalam tiga kategori: positif, netral, dan negatif, berdasarkan nilai compound [3]. Berikut merupakan contoh hasil tahapan *labelling* dapat dilihat pada Gambar 3.

Text	clean	Positive	Negative	Neutral	Compound	Sentiment
Mohon jln.bapak melanggar mau ditilang	Mohon jln.bapak melanggar mau ditilang	0.000	0.000	1.000	0.0000	neutral
Real Madrid's goal has been ruled out for	Real Madrid goal has been ruled out for	0.000	0.074	0.926	-0.1531	negatif
ada yang pernah kena E-tilang? ini	ada yang pernah kena Etilang ini gimana ya	0.000	0.000	1.000	0.0000	neutral
emak siapa nih? bisa lah ya triknya di	emak siapa nih bisa lah ya triknya di pake	0.091	0.000	0.909	0.2023	positif
Semenjak diberlakukannya e-tilang, anak	Semenjak diberlakukannya etilang anak	0.000	0.000	1.000	0.0000	neutral
YUK DOWNLOAD Aplikasi Polri Super	YUK DOWNLOAD Aplikasi Polri Super	0.150	0.000	0.850	0.6981	positif
YUK DOWNLOAD Aplikasi Polri Super	YUK DOWNLOAD Aplikasi Polri Super	0.150	0.000	0.850	0.6981	positif

Gambar 3. Hasil Labelling

3.6. Text Preprocessing

Preprocessing adalah tahap penting dalam proses klasifikasi teks yang mengubah data teks tidak terstruktur menjadi data terstruktur dan bersih. Ini mempermudah proses klasifikasi menggunakan metode naive bayes atau metode klasifikasi teks lainnya dan mempengaruhi kualitas serta akurasi hasil klasifikasi [9].

3.7. Transformation

Pada tahap pertama preprocessing, dilakukan Case Folding untuk menyamaratakan penggunaan huruf kapital pada data input. Hal ini mengubah semua huruf menjadi huruf kecil (lowercase), menghindari perbedaan dalam penggunaan huruf kapital yang dapat mempengaruhi analisis teks [10]. Berikut merupakan contoh hasil dari tahapan *transformation (case folding)* dapat dilihat pada Gambar 4.

index	clean	Sentiment	Tweet
0	Mohon jln.bapak melanggar mau ditilang	neutral	Mohon jln.bapak melanggar mau ditilang sm apa stnk nya
1	Real Madrid's goal has been ruled out for	negatif	Real Madrid goal has been ruled out for
2	ada yang pernah kena Etilang ini gimana ya	neutral	ada yang pernah kena etilang ini gimana ya
3	emak siapa nih bisa lah ya triknya di	positif	emak siapa nih bisa lah ya triknya di

Gambar 4. Hasil Transformation

3.8. Tokenizing

Tokenizing adalah proses memisahkan teks menjadi unit-unit lebih kecil yang disebut token, seperti kata-kata atau kalimat. Tujuannya adalah untuk mempermudah pengelolaan dan analisis data teks [10]. Berikut merupakan contoh hasil *tokenizing* dapat dilihat pada Gambar 5.

Tweet	Token
blue jembatan timbang online etilang user	[blue, jembatan, timbang, online, etilang, use...
cara bayar etilang dan mekanismenya	[cara, bayar, etilang, dan, mekanismenya]
dimana datanya akurat ga coba ambulancenya st	[dimana, datanya, akurat, ga, coba, ambulancen...
hari diresmikan kamera etilang di bandarlampu	[hari, diresmikan, kamera, etilang, di, bandar...
hari kerja kapolri baru salah satunya etilang	[hari, kerja, kapolri, baru, salah, satunya, e...
hari penerapan etilang tercatat ada orang me	[hari, penerapan, etilang, tercatat, ada, oran...
hari tilang elektronik di jalan tol dapat rp	[hari, tilang, elektronik, di, jalan, tol, dap...
kata dia titik lokasi yang dipasang kamera c	[kata, dia, titik, lokasi, yang, dipasang, kam...
lokasi etilang di Palembang rencana selesai g	[lokasi, etilang, di, Palembang, rencana, sele...
marset polda jatang terapan etilang semarang	[marset, polda, jatang, terapan, etilang, sema...

Gambar 5. Hasil Tokenizing

3.9. Filtering

Stop word removal adalah proses menghapus kata-kata tidak penting atau informatif dalam analisis teks atau sentimen. Contoh stop words dalam Bahasa Inggris adalah "and", "or", "this", "the", "is", "in", "of", "a". Dalam Bahasa Indonesia, contoh stop words

adalah "dan", "atau", "ini", "yang", "adalah", "dari", "sebuah", dan lainnya [11]. Berikut merupakan contoh hasil dari tahapan *fitering* dapat dilihat pada Gambar 6.

Token	stopword
[blue, jembatan, timbang, online, etilang, use...	[blue, jembatan, timbang, online, etilang, use...
[cara, bayar, etilang, dan, mekanismenya]	[cara, bayar, etilang, mekanismenya]
[dimana, datanya, akurat, ga, coba, ambulancen...	[datanya, akurat, ga, coba, ambulancenya, stuc...
[hari, diresmikan, kamera, etilang, di, bandar...	[hari, diresmikan, kamera, etilang, bandarlamp...
[hari, kerja, kapolri, baru, salah, satunya, e...	[hari, kerja, kapolri, baru, salah, satunya, e...

Gambar 6. Hasil Filtering

3.10. Stemming

Stemming adalah proses dalam pengolahan teks yang bertujuan untuk menghilangkan imbuhan atau afiks pada kata-kata agar diperoleh kata dasar atau bentuk dasar dari kata tersebut. Imbuhan-imbuhan yang dihilangkan biasanya berupa awalan (prefix) atau akhiran (suffix) pada kata. Proses stemming dilakukan untuk mengurangi variasi bentuk kata dalam teks ke bentuk kata dasarnya [12]. Berikut merupakan contoh hasil *stemming* dapat dilihat pada Gambar 7.

stopword	stemmed
[blue, jembatan, timbang, online, etilang, use...	[blue, jembatan, timbang, online, etilang, use...
[cara, bayar, etilang, mekanismenya]	[cara, bayar, etilang, mekanisme]
[datanya, akurat, ga, coba, ambulancenya, stuc...	[data, akurat, ga, coba, ambulancenya, stuck, ...]
[hari, diresmikan, kamera, etilang, bandarlamp...	[hari, resmi, kamera, etilang, bandarlampung, ...]
[hari, kerja, kapolri, baru, salah, satunya, e...	[hari, kerja, kapolri, baru, salah, satu, etil...

Gambar 7. Hasil Stemming

3.11. TF-IDF

TF-IDF adalah metode yang menghitung bobot kata dalam dokumen dengan menggabungkan Term Frequency dan Inverse Document Frequency. Bobot kata ini menunjukkan pentingnya kata dalam dokumen sambil mempertimbangkan kemunculan kata tersebut dalam seluruh koleksi dokumen. Metode ini sangat berguna dalam analisis teks untuk mengidentifikasi kata-kata yang paling relevan dan informatif dalam suatu dokumen [13]. Berikut merupakan sampel yang digunakan untuk TF-IDF dapat dilihat pada tabel 1.

Tabel 2. Sampel TF-IDF

No	Teks
1.	dapat seluruh layanan polisi genggam mulai urus skck etilang lainlain aplikasi super apps presisi
2.	keluh kesah tiap ngerem auto buta nah gimana dapat
3.	etilang jalan toll max cepat km ga ideal

3.12. Term Frequency

Term Frequency (TF) adalah metode untuk menghitung frekuensi kemunculan suatu kata dalam dokumen. Bobot kata ini adalah jumlah frekuensi

kemunculan kata tersebut dalam dokumen tersebut. Berikut hasil *Term Frequency* dapat dilihat pada tabel 2.

Tabel 3. Hasil *Term Frequency*

TERM	TF		
	d1	d2	d3
dapat	1	1	0
seluruh	1	0	0
layan	1	0	0
polisi	1	0	0
genggam	1	0	0
mulai	1	0	0
urus	1	0	0
skck	1	0	0
etilang	1	0	1
lainlain	1	0	0
aplikasi	1	0	0
super	1	0	0
apps	1	0	0
presisi	1	0	0
keluh	0	1	0
kesah	0	1	0
tiap	0	1	0
ngerem	0	1	0
auto	0	1	0
buta	0	1	0
nah	0	1	0
gimana	0	1	0
jalan	0	0	1
toll	0	0	1
max	0	0	1
cepat	0	0	1
km	0	0	1
ga	0	0	1
ideal	0	0	1

3.13. Inverse Document Frequency

Inverse Document Frequency (IDF) adalah metode untuk mengukur seberapa jarang suatu kata muncul dalam seluruh koleksi dokumen. Bobot kata ini berbanding terbalik dengan frekuensi kemunculannya dalam seluruh dokumen. Semakin jarang kata muncul dalam dokumen, semakin tinggi bobotnya, menunjukkan tingkat keunikannya dalam membedakan dokumen dari yang lain. Rumus IDF dapat dilihat pada persamaan 1:

$IDF = \log(\text{jumlah dokumen} / \text{jumlah frekuensi kata terpilih})$

Berikut merupakan hasil IDF dapat dilihat pada tabel 4.

Tabel 4. Hasil *Inverse Document Frequency*

TERM	DF	IDF
dapat	2	0,176
seluruh	1	0,477
Layan	1	0,477
Polisi	1	0,477
Genggam	1	0,477
Mulai	1	0,477
Urus	1	0,477
Skck	1	0,477

TERM	DF	IDF
Etulang	2	0,176
Lainlain	1	0,477
Aplikasi	1	0,477
Super	1	0,477
Apps	1	0,477
Presisi	1	0,477
Keluh	1	0,477
Kesah	1	0,477
Tiap	1	0,477
Ngerem	1	0,477
Auto	1	0,477
Buta	1	0,477
Nah	1	0,477
Gimana	1	0,477
Jalan	1	0,477
Toll	1	0,477
Max	1	0,477
cepat	1	0,477
km	1	0,477
ga	1	0,477
ideal	1	0,477

Tabel 5. Hasil TF-IDF

TERM	TF IDF		
	d1	d2	d3
dapat	0,176	0,176	0
seluruh	0,477	0	0
layan	0,477	0	0
polisi	0,477	0	0
genggam	0,477	0	0
mulai	0,477	0	0
urus	0,477	0	0
skck	0,477	0	0
etilang	0,176	0	0,176
lainlain	0,477	0	0
aplikasi	0,477	0	0
super	0,477	0	0
apps	0,477	0	0
presisi	0,477	0	0
keluh	0	0,477	0
kesah	0	0,477	0
tiap	0	0,477	0
ngerem	0	0,477	0
auto	0	0,477	0
buta	0	0,477	0
nah	0	0,477	0
gimana	0	0,477	0
jalan	0	0	0,477
toll	0	0	0,477
max	0	0	0,477
cepat	0	0	0,477
km	0	0	0,477
ga	0	0	0,477
ideal	0	0	0,477

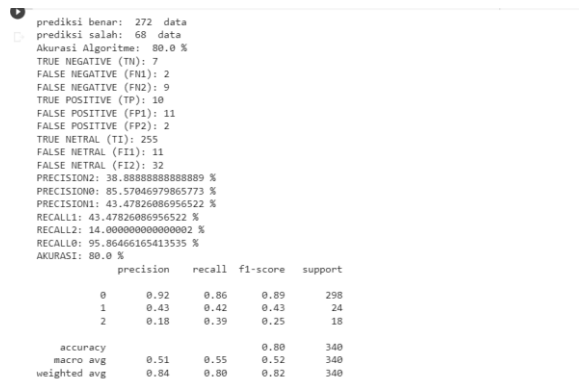
4. HASIL DAN PEMBAHASAN

4.1. Klasifikasi *Naive Bayes*

Naive Bayes merupakan metode pengklasifikasi probabilistik sederhana berdasarkan teorema Bayes. Digunakan dalam analisis teks, pengenalan pola, dan klasifikasi data. Prinsip dasarnya adalah menghitung probabilitas kemunculan kelas berdasarkan atribut

dalam data. Asumsi independensi antar atribut memudahkan perhitungan probabilitas dan implementasi metode ini [8].

4.2. Model Klasifikasi Naive Bayes



```

prediksi benar: 272 data
prediksi salah: 68 data
Akurasi Algoritma: 80.0 %
TRUE NEGATIVE (TN): 7
FALSE NEGATIVE (FN1): 2
FALSE NEGATIVE (FN2): 9
TRUE POSITIVE (TP): 10
FALSE POSITIVE (FP1): 11
FALSE POSITIVE (FP2): 2
TRUE NETRAL (TI): 255
FALSE NETRAL (FI1): 11
FALSE NETRAL (FI2): 32
PRECISION0: 38.88888888888889 %
PRECISION1: 85.57846979865773 %
PRECISION2: 43.47826086956522 %
RECALL1: 43.47826086956522 %
RECALL2: 14.000000000000002 %
RECALL0: 95.86466165413535 %
AKURASI: 80.0 %
  
```

	precision	recall	f1-score	support
0	0.92	0.86	0.89	298
1	0.43	0.42	0.43	24
2	0.18	0.39	0.25	18
accuracy			0.80	340
macro avg	0.51	0.55	0.52	340
weighted avg	0.84	0.80	0.82	340

Gambar 8. Hasil Klasifikasi Naive Bayes

4.3. Evaluasi (Confusion Matrix)

Confusion Matrix adalah tabel yang membandingkan hasil klasifikasi model dengan kelas sebenarnya. Dari sini, kita dapat menghitung metrik evaluasi kinerja model seperti akurasi, presisi, dan recall. Akurasi mengukur ketepatan klasifikasi keseluruhan, presisi mengukur identifikasi data positif, dan recall mengukur kemampuan menemukan semua data positif [8].

1. Perhitungan Akurasi

$$\text{Akurasi} = \frac{TP + TN + TI}{\text{Jumlah Data}} \times 100\% = \frac{10 + 7 + 255}{339} \times 100\% = \frac{272}{339} \times 100\% = 0,8 \times 100\% = 80\%$$

Rumus 1. Perhitungan Akurasi

2. Perhitungan Presisi Positif

$$\begin{aligned} \text{Presisi Positif} &= \frac{TP}{TP + FP1 + FP2} \times 100\% \\ &= \frac{10}{10 + 11 + 2} \times 100\% \\ &= \frac{10}{23} \times 100\% \\ &= 0,43 \times 100\% = 43\% \end{aligned}$$

Rumus 2. Perhitungan Presisi Positif

3. Perhitungan Presisi Negatif

$$\begin{aligned} \text{Presisi Negatif} &= \frac{TN}{TN + FN1 + FN2} \times 100\% \\ &= \frac{7}{7 + 2 + 9} \times 100\% \\ &= \frac{7}{18} \times 100\% \end{aligned}$$

$$= 0,38 \times 100\% = 38\%$$

Rumus 3. Perhitungan Presisi Negatif

4. Perhitungan Presisi Netral

$$\begin{aligned} \text{Presisi Netral} &= \frac{TI}{TI + FI1 + FI2} \times 100\% \\ &= \frac{255}{255 + 11 + 32} \times 100\% \\ &= \frac{255}{298} \times 100\% \\ &= 0,85 \times 100\% = 85\% \end{aligned}$$

Rumus 4. Perhitungan Presisi Netral

5. Perhitungan Recall Positif

$$\begin{aligned} \text{Recall Positif} &= \frac{TP}{FN1 + FI1 + TP} \times 100\% \\ &= \frac{10}{2 + 11 + 10} \times 100\% = \frac{10}{23} \times 100\% \\ &= 0,43 \times 100\% = 43\% \end{aligned}$$

Rumus 5. Perhitungan Recall Positif

6. Perhitungan Recall Negatif

$$\begin{aligned} \text{Recall Negatif} &= \frac{TN}{FP1 + FI2 + TN} \times 100\% \\ &= \frac{7}{11 + 32 + 7} \times 100\% = \frac{7}{50} \times 100\% \\ &= 0,14 \times 100\% = 14\% \end{aligned}$$

Rumus 6. Perhitungan Recall Negatif

7. Perhitungan Recall Netral

$$\begin{aligned} \text{Recall Netral} &= \frac{TI}{FN2 + FP2 + TI} \times 100\% \\ &= \frac{255}{9 + 2 + 255} \times 100\% \\ &= \frac{255}{266} \times 100\% \\ &= 0,95 \times 100\% = 95\% \end{aligned}$$

Rumus 7. Perhitungan Recall Netral

4.4. Word Cloud

Word cloud adalah metode efektif untuk memvisualisasikan informasi dari dokumen dalam kategori negatif, positif, dan netral tentang sistem e-tilang. Word cloud digunakan untuk menampilkan kata-kata yang sering muncul dalam penelitian ini. Kata-kata yang muncul lebih sering ditampilkan lebih besar dan mencolok, memberikan gambaran visual tentang frekuensi kata-kata yang relevan dalam setiap kategori [8]. Berikut merupakan hasil dari tahapan *word cloud* dapat dilihat pada Gambar 9.



Gambar 9. Hasil *Word Cloud*

5. KESIMPULAN DAN SARAN

Hasil penelitian analisis sentimen terhadap sistem e-tilang menunjukkan tanggapan masyarakat cenderung netral. Setelah melalui proses Confusion Matrix didapatkan presisi netral 85% dan recall netral 95%, tetapi perlu ditingkatkan pada kelas positif dan negatif. Akurasi keseluruhan mencapai 80%. Analisis ini memberikan informasi penting bagi pemerintah dalam memahami respons masyarakat terhadap sistem e-tilang.

DAFTAR PUSTAKA

-
- Gambar 9. Hasil Word Cloud
- ## 5. KESIMPULAN DAN SARAN
- Hasil penelitian analisis sentimen terhadap sistem e-tilang menunjukkan tanggapan masyarakat cenderung netral. Setelah melalui proses Confusion Matrix didapatkan presisi netral 85% dan recall netral 95%, tetapi perlu ditingkatkan pada kelas positif dan negatif. Akurasi keseluruhan mencapai 80%. Analisis ini memberikan informasi penting bagi pemerintah dalam memahami respons masyarakat terhadap sistem e-tilang.
- ## DAFTAR PUSTAKA
- [1] Apriliana Zunia Lutfina, “EFEKTIVITAS PENGGUNAAN E-TILANG TERHADAP PELANGGARAN LALU LINTAS DI POLRES MAGELANG,” *J. Komun. Huk. Univ. Pendidik. Ganesha*, 2019.
 - [2] K. Manso, F. D. J. Lengkong, and Y. Londa, “IMPLEMENTASI PROGRAM ELEKTRONIK TILANG (E-TILANG) PADA DIREKTORAT LALU LINTAS (DITLANTAS) POLDA SULAWESI UTARA,” 2022.
 - [3] A. Retno *et al.*, “Sistem Koreksi Kesalahan Pengetikan Kata Kunci dalam Pencarian Artikel Menggunakan Algoritma Jaro-Winkler,” *Semin. Inform. Apl. Polinema*, 2020.
 - [4] I. Gunaawan and Y. Fernando, “SISTEM PAKAR DIAGNOSA PENYAKIT KULIT PADA KUCING MENGGUNAKAN METODE NAIVE BAYES BERBASIS WEB,” *J. Inform. dan Rekayasa Perangkat Lunak*, vol. 2, no. 2, pp. 239–247, 2021, [Online]. Available: <http://jim.teknokrat.ac.id/index.php/informatika>
 - [5] T. Ernayanti, M. Mustafid, A. Rusgiyono, and A. R. Hakim, “PENGGUNAAN SELEKSI FITUR CHI-SQUARE DAN ALGORITMA MULTINOMIAL NAÏVE BAYES UNTUK ANALISIS SENTIMEN PELANGGAN TOKOPEDIA,” *J. Gaussian*, vol. 11, no. 4, pp. 562–571, Feb. 2023, doi: 10.14710/j.gauss.11.4.562-571.
 - [6] M. Arifin, K. Prasetyo, J. Vivaldi, A. Sander, K. Khothimah, and R. Ardhana, “Tingkat Kesadaran Masyarakat Terhadap Penerapan Sistem E-Tilang Kota Surabaya,” *Student J. Public Manag.*, 2020.
 - [7] N. M. A. J. Astari, Dewa Gede Hendra Divayana, and Gede Indrawan, “Analisis Sentimen Dokumen Twitter Mengenai Dampak Virus Corona Menggunakan Metode Naive Bayes Classifier,” *J. Sist. dan Inform.*, vol. 15, no. 1, pp. 27–29, Nov. 2020, doi: 10.30864/jsi.v15i1.332.
 - [8] E. Manalu, F. A. Sianturi, and M. R. Manalu, “PENERAPAN ALGORITMA NAIVE BAYES UNTUK MEMPREDIKSJUMLAH PRODUKSI BARANG BERDASARKAN DATAPERSEDIAAN DAN JUMLAH PEMESANAN PADACV. PAPADAN MAMA PASTRIES,” *J. Mantik Penusa*, vol. 1, no. 2, p. 1, 2021.
 - [9] H. Ma’rifah, A. Prasetya Wibawa, and M. I. Akbar, “Sains, Aplikasi, Komputasi dan Teknologi Informasi Klasifikasi artikel ilmiah dengan berbagai skenario preprocessing,” *Sains, Apl. Komputasi dan Teknol. Inf.*, vol. 2, no. 2, p. 70, 2020.
 - [10] E. B. Santoso and A. Nugroho, “Analisis Sentimen Calon Presiden Indonesia 2019 Berdasarkan Komentar Publik Di Facebook,” *Eksplora Inform.*, vol. 9, no. 1, pp. 60–69, Sep. 2019, doi: 10.30864/eksplora.v9i1.254.
 - [11] S. Ramadhani, D. Azzahra, and T. Z., “Comparison of K-Means and K-Medoids Algorithms in Text Mining based on Davies Bouldin Index Testing for Classification of Student’s Thesis,” *Digit. Zo. J. Teknol. Inf. dan Komun.*, vol. 13, no. 1, pp. 24–33, May 2022, doi: 10.31849/digitalzone.v13i1.9292.
 - [12] A. R. Hidayati, A. S. Fitriani, M. A. Rosid, F. Sains, and D. Teknologi, “Analisa Sentimen Pemilu 2019 Pada Judul Berita Online Menggunakan Metode Logistic Regression,” *J. Penerapan Sist. Inf.*, vol. 4, no. 2, pp. 298–305, 2023.
 - [13] D. Atika, A. Ari Aldino, S. Informasi, J. Pagar Alam No, L. Ratu, and K. Kedaton, “TERM FREQUENCY-INVERSE DOCUMENT FREQUENCY SUPPORT VECTOR MACHINE UNTUK ANALISIS SENTIMEN OPINI MASYARAKAT TERHADAP TEKATAN MENTAL PADA MEDIA SOSIAL TWITTER,” 2022. [Online]. Available: <http://jim.teknokrat.ac.id/index.php/JTSI>