



Analisis Sentimen Warganet Terhadap Keberadaan Juru Parkir Liar Menggunakan Metode Naive Bayes Classifier

Avis Tantra Mukti, Firman Noor Hasan*

Fakultas Teknologi Industri dan Informatika, Teknik Informatika, Universitas Muhammadiyah Prof. Dr. Hamka, Jakarta, Indonesia

Email: ¹avistantra01@gmail.com, ^{2,*}firman.noorhasan@uhamka.ac.id

Email Penulis Korespondensi: firman.noorhasan@uhamka.ac.id

Abstrak—Masih tingginya angka pengangguran di Indonesia menyebabkan banyak sekali dampak, salah satunya adalah munculnya juru parkir liar. Kondisi tersebut terus didukung dengan terciptanya area parkir para pelaku usaha untuk para pengunjungnya. Area parkir yang disediakan sering kali gratis, bahkan memiliki tulisan yang menyatakannya. Akan tetapi, ada saja oknum yang memanfaatkan lahan parkir gratis tersebut untuk mendapatkan penghasilan. Banyak sekali sentimen dari warganet terhadap fenomena juru parkir liar di media sosial. Oleh sebab itu dalam penelitian ini digunakan sebuah analisis tanggapan berupa sentimen warganet terhadap juru parkir liar di media sosial X dengan menggunakan Naive Bayes. Tujuan utama penelitian ini adalah memahami perasaan masyarakat terhadap keberadaan juru parkir liar yang beroperasi di area parkir. Dataset yang digunakan dalam analisis ini sebanyak 905 yang diambil dari media sosial X dan dianalisis menggunakan algoritma Naive Bayes dengan melakukan klasifikasi sentimen ke dalam kategori positif, netral, dan negatif terhadap keberadaan juru parkir liar yang beroperasi. Hasil dari penelitian ini warganet merasa sangat kesal, marah, dan terganggu dengan keberadaan juru parkir liar yang beroperasi. Hal ini dibuktikan dengan hasil sentimen negatif yang mendominasi sebesar 93% dari keseluruhan data atau sebanyak 841 sentimen negatif terhadap fenomena ini.

Kata Kunci: Analisis Sentimen; Juru Parkir Liar; Media Sosial X; Tweet; Naive Bayes Classifier

Abstract—The still high rate of poverty in Indonesia causes many impacts, one of which is the emergence of illegal parking attendants. This condition continues to be supported by the creation of a parking area for business actors for their visitors. The parking areas provided are often free, and even have signs saying so. However, there are individuals who use the free parking space to earn income. There are many netizens' sentiments regarding the phenomenon of lying parking attendants on social media. Therefore, in this research, an analysis was used in the form of netizen sentiment towards illegal parking attendants on social media X using Naive Bayes. The main objective of this research is to understand the public's feelings towards the existence of illegal parking attendants operating in the parking area. The dataset used in this analysis was 905 taken from social media. The results of this research netizens felt very annoyed, angry and disturbed by the presence of illegal parking attendants operating. This is proven by the results of negative sentiment which dominates 93% of the total data or as many as 841 negative sentiments regarding this phenomenon.

Keywords: Sentiment Analysis; Illegal Parking Officer; Social Media X; Tweets; Naive Bayes Classifier

1. PENDAHULUAN

Di era digital ini, peran media sosial sangat digemari karena dapat digunakan untuk berbagi pengalaman dan pendapat mengenai beragam topik secara terbuka tanpa adanya batasan termasuk keresahan yang dialami oleh penggunanya [1]. Seiring berkembangnya teknologi, kini aplikasi didukung dengan adanya versi mobile sehingga berbagi pengalaman penggunaan lebih mudah dan fleksibel untuk dilakukan di mana pun kapan pun [2] asal terhubung dengan internet [3]. Sementara itu, manusia sebagai makhluk ekonomi yang perlu melakukan transaksi tidak lepas dari yang namanya praktik jual beli. Pengguna media sosial khususnya X sering kali mengunggah pengalaman keseharian dan opininya [4]. Tak terkecuali pengalaman jual beli, khususnya perihal parkir. Perkara parkir pada kegiatan jual beli sering kali dibagikan melalui media sosial X terlepas pengalaman menyenangkan maupun pengalaman buruk.

Sesuai peraturan, lahan parkir adalah suatu keharusan bagi pelaku usaha sesuai undang-undang yang berlaku dalam peraturan presiden nomor 112 tahun 2007 [5]. Hal ini berguna untuk mengantisipasi terjadinya kemacetan karena pembeli atau pengunjung yang datang. Setiap pelaku usaha memiliki hak masing-masing untuk membuat kebijakan apakah area parkir yang disediakan gratis atau berbayar. Akan tetapi tidak jarang area parkir gratis yang disediakan oleh pelaku usaha tidak sesuai dengan kenyataannya sehingga membuat pembeli atau pengunjung merasa kurang nyaman. Hal tersebut dikarenakan maraknya muncul oknum juru parkir liar yang memanfaatkan situasi ini [6] ditambah keberadaannya diperkirakan dilindungi oleh pihak lain yang sama-sama mengambil keuntungan [7]. Juru parkir liar kerap melakukan pungutan [8] kepada pengguna parkir ketika mereka akan pergi meninggalkan area parkir. Pungutan liar sendiri telah dilarang dalam undang-undangan dan dapat dikenai pidana [9]. Maraknya oknum yang bermunculan juga dikarenakan masih tingginya jumlah pengangguran di Indonesia yang mencapai ratusan juta jiwa [10]. Hal tersebut terjadi karena pertumbuhan tenaga kerja tidak sejalan dengan tumbuhnya lapangan pekerjaan [11].

Banyaknya pengguna media sosial X yang memiliki pengalaman tentang keberadaan juru parkir liar, menghasilkan banyak sekali tweet di media sosial X. Dalam konteks ini, analisis sentimen sangat berguna sebagai alat untuk memahami pemikiran warganet terhadap munculnya juru parkir liar. Hal ini juga bisa sekaligus mengetahui respon warganet terhadap kinerja aparat penegak hukum untuk mengatasi kegiatan ilegal ini. Dengan



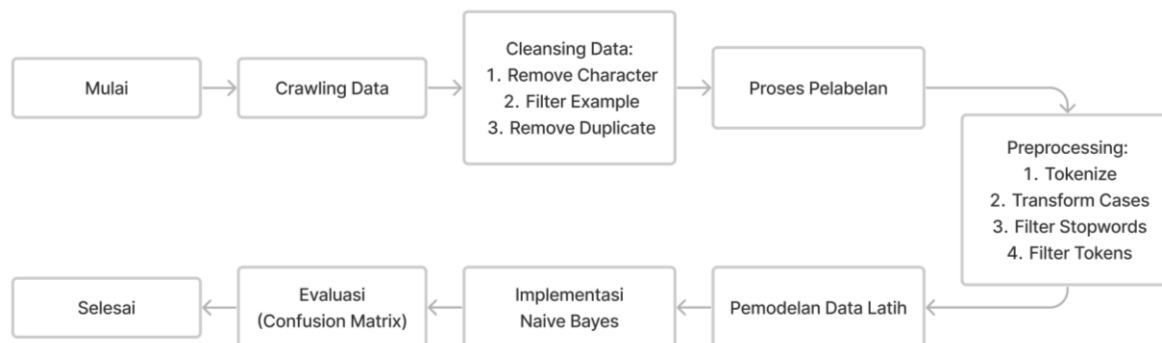
menganalisis emosi, pendapat, dan sikap yang diungkapkan oleh pengguna media sosial, maka analisis sentimen memungkinkan bisnis dan organisasi untuk mengumpulkan informasi dan mengambil keputusan untuk strategi bisnis selanjutnya [12].

Penelitian terdahulu terkait analisis sentimen telah dilakukan sebelumnya. Diantaranya adalah penelitian untuk menganalisis sentimen masyarakat terhadap dampak pandemi akibat virus corona dengan menggunakan metode Naïve Bayes Classifier dan Support Vector Machine. Penelitian ini menggunakan data sebanyak 1104 yang diambil dari media sosial Twitter. Hasil penelitian tersebut menunjukkan tingkat akurasi sebesar 81,07% menggunakan Naïve Bayes Classification dan 79,96% menggunakan Support Vector Machine [13]. Penelitian selanjutnya membahas sentimen terhadap program kampus merdeka menggunakan algoritma Naïve Bayes. Pada penelitian ini menggunakan data sebanyak 501 opini publik yang dari media sosial Twitter. Hasil penelitian tersebut menunjukkan tingkat akurasi sebesar 57% dengan opini positif sebanyak 272 dan opini negatif sebanyak 229 [14]. Penelitian analisis sentimen lainnya membahas tentang ulasan pelanggan pada Aplikasi Fore Coffe yang tersaji pada Google Play Store menggunakan metode Naïve Bayes, juga mendapatkan hasil akurasi yang tinggi sebesar 74,28% dari 1801 data yang diproses [15]. Penelitian lainnya adalah menganalisis sentimen yang membahas tentang pembubaran konser NCT 127 sebanyak 2451 data menggunakan metode Naïve Bayes dengan hasil penelitian mendapatkan akurasi sebesar 82,01% [16]. Penelitian lainnya yang menggunakan metode Naive Bayes adalah tentang keberadaan ChatGPT di Indonesia, juga menunjukkan bahwa hasil akurasi dari metode ini adalah sebesar 77,33% dari 754 data dengan hasil sentimen 436 negatif dan 318 positif [17].

Berdasarkan penelitian terdahulu yang telah dilakukan dan fenomena yang terjadi, belum pernah ada analisis sentimen yang berkaitan tentang keberadaan juru parkir liar di masyarakat. Oleh sebab itu, penelitian ini bertujuan untuk menganalisis opini masyarakat Indonesia mengenai keberadaan juru parkir liar dengan menggunakan metode Naïve Bayes sebagai algoritma untuk klasifikasi data. Hal ini bertujuan untuk mengetahui respon masyarakat dengan adanya juru parkir liar sekaligus menjadi bahan evaluasi bagi penegak hukum untuk mengatasi kegiatan ilegal tersebut. Selain itu juga dapat menjadi bahan pertimbangan untuk mencari solusi peraturan bagi pemerintah dalam mengatasi praktik ilegal ini. Sumber dataset yang digunakan berasal dari tweet media sosial X dengan menggunakan kata kunci “tukang parkir liar” dalam rentang waktu 5 Oktober 2023. Total tweet yang berhasil dikumpulkan sebanyak 928 tweet. Melalui proses analisis, penelitian ini mengklasifikasikan data menjadi tiga jenis sentimen, yaitu sentimen positif, netral, dan negatif.

2. METODOLOGI PENELITIAN

Dalam pengembangan penelitian ini, digunakan metode Naïve Bayes untuk melakukan analisis sentimen. Metode Naïve Bayes adalah klasifikasi dengan menghitung probabilitas dengan teorema Bayes dengan kombinasi nilai frekuensi dari data [18] dengan memprediksi peluang di masa depan berdasarkan pengalaman sebelumnya [19]. Menggunakan algoritma sederhana dan memiliki akurasi yang tinggi merupakan kelebihan dari Naive Bayes [20][21]. Bahkan metode ini lebih unggul jika dibandingkan dengan algoritma Support Vector Machine dan K-Nearest Neighbor [22]. Terdapat diagram alir yang menggambarkan langkah demi langkah dalam penelitian ini agar terstruktur dan dapat terukur. Diagram tersebut sesuai yang tertera pada gambar 1.



Gambar 1. Diagram Alir Tahapan Penelitian

Untuk memulai penelitian, tahap pertama adalah crawling data dari internet yang berfokus pada media sosial X. Crawling data adalah pengambilan data pada internet. Pengambilan data melalui media sosial Twitter karena X adalah salah satu media sosial yang mengizinkan pengguna untuk berbagi foto, video, hingga pendapat dengan beropini secara bebas [23]. Pengambilan data di media sosial X melalui Google Colab menggunakan bahasa Python dengan output dalam bentuk file CSV. Setelah itu, data yang telah berhasil diambil akan dilakukan pemrosesan berupa preprocessing. Pada tahap ini data akan dirapikan dengan cara menghapus bagian data yang berupa URL, mention, hastag, karakter karakter yang tidak diperlukan ([!~?.,:."#*%@&()-]), data kosong, data duplikat, dan whitespace. Tahap ini bertujuan untuk memudahkan peneliti dalam melakukan pelabelan sentimen.



Setelah dataset telah rapi dari karakter karakter yang tidak diperlukan, maka siap dilanjutkan ke tahap labeling. Pada tahap ini data akan diberikan label sentimen apakah data mengandung sentimen positif, negatif, ataupun netral dengan menggunakan Microsoft Excel secara manual. Pelabelan ini dilakukan hanya sebagian saja dari keseluruhan data untuk disebut sebagai data latih. Sebanyak 543 data akan dilabeli dan 362 dibiarkan saja dengan pembagian data menggunakan perbandingan 60:40 dari keseluruhan data. Nantinya data latih inilah yang dijadikan sebagai contoh melalui metode Naïve Bayes untuk menentukan data yang belum dilabeli atau yang disebut data uji. Setelah proses Naïve Bayes selesai dalam memproses data uji, maka data latih yang telah dilabeli di awal akan digabung dengan hasil data uji yang telah didapatkan. Setelah data latih dan data uji digabung, maka dilakukan evaluasi untuk mengukur akurasi dari hasil proses yang telah dilakukan. Metode yang digunakan untuk menghitung tingkat akurasi adalah confusion matrix dengan 3 parameter, yakni accuracy, precision, dan recall.

3. HASIL DAN PEMBAHASAN

3.1 Data Penelitian

Penelitian ini menggunakan dataset yang berasal dari media sosial X. Dataset diperoleh menggunakan bahasa Python melalui Google Colab dengan proses crawling. Total data yang terkumpul sebanyak 928 tweet pada tanggal 5 Oktober 2023. Proses crawling sesuai pada gambar 2 berikut ini.



Gambar 2. Proses Crawling Data

Peneliti menggunakan kata kunci "tukang parkir liar" untuk menyaring tweet yang akan diambil dari media sosial X sesuai yang terdapat pada gambar 3. Setelah data berhasil diambil, kemudian data disimpan sementara pada file yang bernama jukir_liar dalam format CSV. Setelah proses crawling selesai, kemudian file CSV akan diunduh dan akan dilanjutkan kepada proses preprocessing.

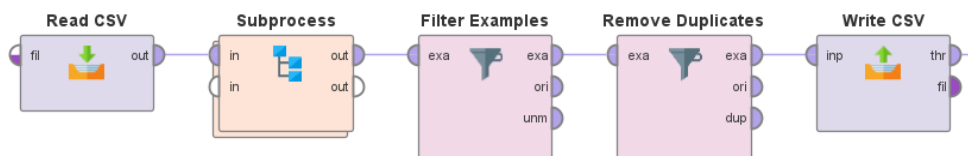
```
# Crawl Data
filename = 'jukir_liar.csv'
search_keyword = 'tukang parkir liar'
limit = 999
```

Gambar 3. Crawling Data X Menggunakan kata kunci "Tukang Parkir Liar"

Variabel "filename" berfungsi untuk memberikan nama file yang akan disimpan dari hasil crawling yang telah dilakukan. Variabel "search_keyword" berfungsi untuk mengatur kata kunci pencarian yang akan digunakan untuk mencari data terkait juru parkir liar. Variabel "limit" berfungsi untuk mengatur banyaknya data yang akan diambil dari internet. Dan baris terakhir adalah kode yang berfungsi untuk memasukkan token personal agar dapat terhubung ke media sosial X.

3.2 Cleansing

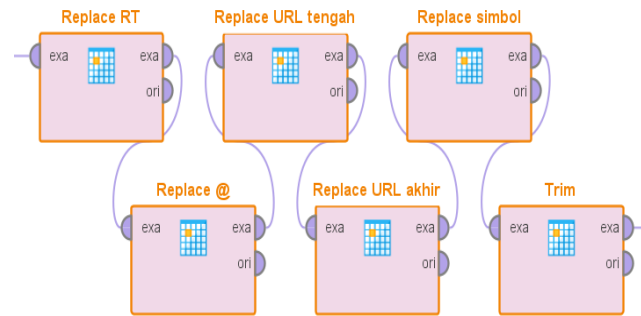
Tahap cleansing berfungsi menyeleksi data agar data menjadi beraturan supaya memudahkan dalam proses penelitian di tahap selanjutnya. Data yang telah diambil kemudian akan dibersihkan sesuai kebutuhan [24]. Tahap ini menggunakan bantuan aplikasi RapidMiner. Beberapa operator digunakan pada tahap ini seperti yang disajikan pada gambar 4.



Gambar 4. Desain Operator pada Tahap Cleansing

Gambar diatas merupakan tahapan pada cleansing dengan penjelasan sebagai berikut:

- a. Read CSV
Read CSV merupakan operator yang berfungsi untuk membaca file CSV. Pada tahap ini digunakan file CSV yang berisi data tweet telah diunduh dari hasil crawling data di internet.
- b. Subprocess
Subprocess merupakan operator yang didalamnya dapat berisi kumpulan operator.



Gambar 5. Desain Beberapa Operator dalam Operator Subprocess

Pada penelitian ini, Operator subprocess diisi dengan operator “Replace” dan “Trim”. Operator Replace untuk menghapus dan membersihkan data dari teks dan karakter yang tidak diperlukan, yakni, URL, mention, hastag, dan karakter karakter yang tidak diperlukan ([!~?.,:."#*%@%\$&()-]). Sedangkan operator Trim untuk menghapus spasi berlebih karena adanya proses berupa replace sehingga terjadinya perubahan pada data. Desain operator yang digunakan tertera pada gambar 5. Perbandingan data sebelum dan sesudah dilakukan proses cleansing ditampilkan pada tabel 1.

Tabel 1. Perbandingan Hasil Cleansing

Sebelum	Sesudah
<p>@infoXwor_ Bisa bilang 'gak' sama tukang parkir liar dan pengamen ga niat. Bahkan kadang ga pake bilang. Langsung ngeloyor ato cuekin aja. Ngapain sungkan sama yg tidak tertib.</p> <p>tolong dong ini kalau mau meniadakan belanja online (tiktokshop dll), itu tukang parkir yg gajelas ditiadakan juga.. malesnya belanja offline ya itu krn tukang parkir liar yg gaje itu. mana biaya parkirnya mahal bgtt ðŸŸ²</p> <p>@tanyarlfe Tp tergantung. Klo kang parkir liar dan mereka cm mintain duit ya kesel, tp kalo tukang parkir beneran ya engga</p> <p>di dunia nyata: tukang parkir liar di internet: affiliator shopee pekerjaan paling dicintai di muka bumi.</p> <p>GUE MELIHAT DIRI GW DI BAYU..... Believe or not, dikantor kan parkir di parkir liar gt. Kalo ngutang boleh krn gw bestie sama yg jaga parkirnya. Dia selalu jagain motor gw yang liputan sampe tengah malem biar gw ga bingung nyarinya. BEST TUKANG PARKIR EVER</p>	<p>Bisa bilang gak sama tukang parkir liar dan pengamen ga niat Bahkan kadang ga pake bilang Langsung ngeloyor ato cuekin aja Ngapain sungkan sama yg tidak tertib</p> <p>tolong dong ini kalau mau meniadakan belanja online tiktokshop dll itu tukang parkir yg gajelas ditiadakan juga malesnya belanja offline ya itu krn tukang parkir liar yg gaje itu mana biaya parkirnya mahal bgtt</p> <p>Tp tergantung Klo kang parkir liar dan mereka cm mintain duit ya kesel tp kalo tukang parkir beneran ya engga</p> <p>di dunia nyata tukang parkir liar di internet affiliator shopee pekerjaan paling dicintai di muka bumi</p> <p>GUE MELIHAT DIRI GW DI BAYU Believe or not dikantor kan parkir di parkir liar gt Kalo ngutang boleh krn gw bestie sama yg jaga parkirnya Dia selalu jagain motor gw yang liputan sampe tengah malem biar gw ga bingung nyarinya BEST TUKANG PARKIR EVER</p>

- c. Filter Example
Kegagalan dan kesalahan sangat mungkin terjadi pada saat crawling data, maka dari itu operator ini perlu digunakan untuk akan menyaring dan memilih data yang tidak hilang untuk diolah.
- d. Remove Duplicate
Operator ini berfungsi untuk menghapus data yang sama agar menjadi unik (satu) agar tidak ada duplikasi pada data yang akan diolah.
- e. Write CSV
Operator ini berfungsi untuk menuliskan kembali hasil dari data yang telah diproses menjadi file dengan format CSV.

3.3 Proses Pelabelan

Dataset yang telah melalui cleansing selanjutnya masuk ke tahap pelabelan [25]. Pada tahap ini, semua data yang telah disaring dan dibersihkan dibagi menjadi dua yakni data latih dan data uji. Pada data latih diberikan pelabelan sebagai contoh untuk mengolah data uji nantinya. Dari 928 data yang berhasil dikumpulkan dari internet, hanya sebanyak 905 data yang akan digunakan dalam penelitian ini. Data yang tidak relevan akan dibuang agar penelitian dapat semakin akurat. Peneliti membagi data dengan perbandingan 60:40. Dari 905 data, sebanyak 543 data dijadikan sebagai data latih dan 362 menjadi data uji. Pada data latih, diberikan label positif jika data mengandung opini positif, label negatif jika data mengandung opini negatif, dan label netral jika data mengandung opini netral [26] sesuai yang ditunjukkan pada tabel 2.

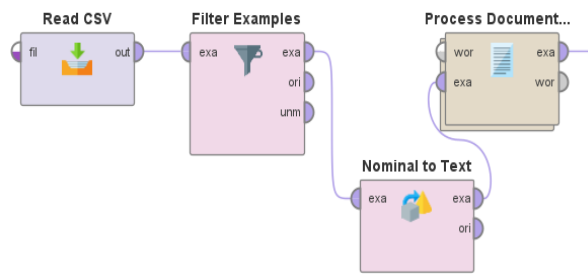


Tabel 2. Contoh Pelabelan Sentimen

Opini	Sentimen
Bisa bilang gak sama tukang parkir liar dan pengamen ga niat Bahkan kadang ga pake bilang Langsung ngeloyor ato cuekin aja Ngapain sungkan sama yg tidak tertib	Negatif
tolong dong ini kalau mau meniadakan belanja online tiktokshop dll itu tukang parkir yg gajelas ditiadakan juga malesnya belanja offline ya itu krn tukang parkir liar yg gaje itu mana biaya parkirnya mahal bgtt	Negatif
Tp tergantung Klo kang parkir liar dan mereka cm mintain duit ya kesel tp kalo tukang parkir beneran ya engga	Netral
di dunia nyata tukang parkir liar di internet affiliator shopee pekerjaan paling dicintai di muka bumi	Positif
GUE MELIHAT DIRI GW DI BAYU Believe or not dikantor kan parkir di parkir liar gt Kalo ngutang boleh krn gw bestie sama yg jaga parkirnya Dia selalu jagain motor gw yang liputan sampe tengah malem biar gw ga bingung nyarinya BEST TUKANG PARKIR EVER	Positif

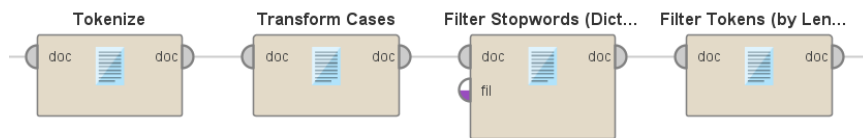
3.4 Preprocessing

Proses selanjutnya adalah menjadikan sisa datanya sebanyak 362 sebagai data uji yang akan diproses pada RapidMiner menggunakan metode Naïve Bayes.



Gambar 6. Import File Data Latih

Pertama-tama menggunakan operator “Read CSV” untuk membaca file CSV yang telah disiapkan. File CSV yang digunakan adalah file yang berisi data latih. Kemudian operator dihubungkan dengan operator “Filter Example”. Fungsi dari operator ini adalah untuk menyaring data sesuai kebutuhan. Untuk itu, digunakan kondisi menyaring sentimen yang tidak hilang dengan menggunakan opsi “is not missing”. Dari operator “Filter Example” selanjutnya akan dihubungkan dengan operator “Process Document from Data”, akan tetapi perlu diselipkan operator “Nominal to Text” terlebih dahulu untuk mengubah nilai nominal menjadi Text agar operator “Process Document from Data” dapat membaca data dan berfungsi. Di dalam operator “Process Document from Data” terdapat beberapa operator yang digunakan seperti yang disajikan pada gambar 7.



Gambar 7. Operator Yang Digunakan Dalam Preprocessing

- a. Tokenize
Operator ini berfungsi untuk memisahkan kata-kata dalam kalimat menjadi individual seperti yang tertera pada tabel 3.

Tabel 3. Perbandingan Hasil Operator Tokenize

Tanpa Tokenize	Dengan Tokenize
Bisa bilang gak sama tukang parkir liar dan pengamen ga niat Bahkan kadang ga pake bilang Langsung ngeloyor ato cuekin aja Ngapain sungkan sama yg tidak tertib	‘Bisa’, ‘bilang’, ‘gak’, ‘sama’, ‘tukang’, ‘parkir’, ‘liar’, ‘dan’, ‘pengamen’, ‘ga’, ‘niat’, ‘Bahkan’, ‘kadang’, ‘ga’, ‘pake’, ‘bilang’, ‘Langsung’, ‘ngeloyor’, ‘ato’, ‘cuekin’, ‘aja’, ‘Ngapain’, ‘sungkan’, ‘sama’, ‘yg’, ‘tidak’, ‘tertib’

- b. Transform Cases
Operator ini berfungsi untuk menyeragamkan kapitalisasi pada setiap kata dalam data. Untuk ini digunakan pilihan “lower case” agar seluruh data seragam menjadi huruf kecil.



Tabel 4. Perbandingan Hasil Operator Transform Case

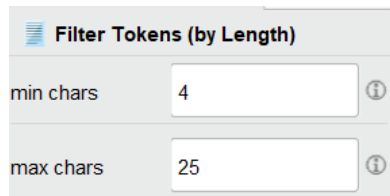
Tanpa Transform Cases	Dengan Transform Cases
Bisa bilang gak sama tukang parkir liar dan pengamen ga niat Bahkan kadang ga pake bilang Langsung ngeloyor ato cuekin aja Ngapain sungkan sama yg tidak tertib	bisa bilang gak sama tukang parkir liar dan pengamen ga niat bahkan kadang ga pake bilang langsung ngeloyor ato cuekin aja ngapain sungkan sama yg tidak tertib

c. Filter Stopword (Dictionary)

Operator ini berfungsi untuk menyaring dan menghapus kata kata pada data yang terkandung dalam file stopwords. File stopwords sendiri bersumber dari kajian yang terdapat pada referensi [27].

d. Filter Tokens (by Length)

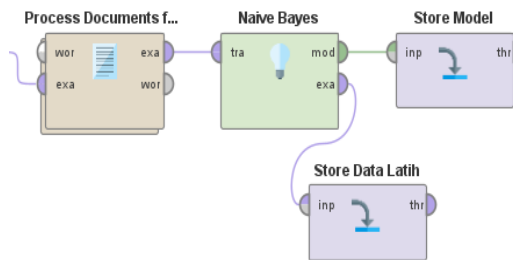
Operator ini berfungsi untuk menyaring dan menghapus setiap data sesuai dari batas tertentu berdasarkan panjang. Batas yang digunakan adalah minimal karakter 4 dan maksimal karakter 25 sesuai pada gambar 8.



Gambar 8. Parameter Panjang Tokens Yang Diatur

3.5 Pemodelan Data Latih

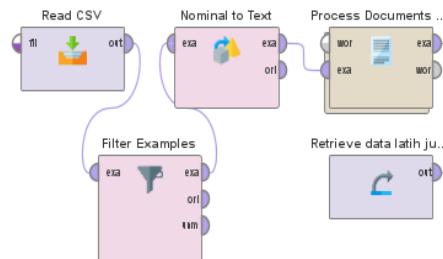
Selanjutnya pada operator “Process Document from Data” dihubungkan dengan operator “Naïve Bayes”. Setelah itu, operator “Naïve Bayes” dihubungkan lagi dengan 2 operator “Store”. Operator “Store” yang pertama berfungsi untuk menyimpan hasil proses dari operator “Naïve Bayes” yang telah dinamai ulang menjadi “Store Model”. Sementara operator “Store” lainnya dinamai dengan “Store Data Latih” untuk menyimpan hasil dari data latih sesuai dengan gambar 9.



Gambar 9. Operator Yang Digunakan Dalam Pemodelan Data Latih

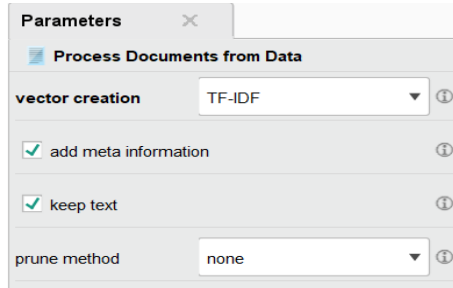
3.6 Implementasi Naïve Bayes

Setelah pembuatan data latih berhasil, selanjutnya adalah Implementasi Naïve Bayes. Operator RapidMiner yang digunakan kali ini dapat dilihat pada gambar 10.



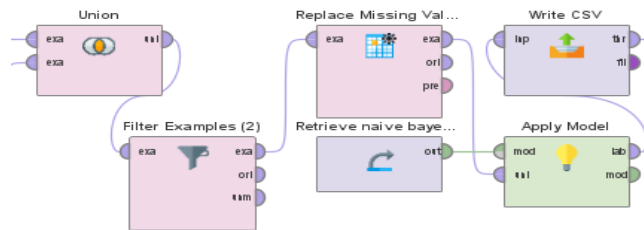
Gambar 10. Operator untuk import Data Latih

Implementasi penggunaan metode Naïve Bayes diawali dengan menggunakan operator “Read CSV” untuk membaca file CSV yang telah disiapkan, kemudian hubungkan dengan operator “Filter Example”. Operator ini diatur pada pilihan “is missing” untuk nantinya sentimen yang masih kosong atau belum dilabeli pada data latih yang akan diambil dan diproses. Selanjutnya agar data dapat diproses, diperlukan operator “Nominal to Text” agar operator “Process Document from Data” dapat membaca dan memproses data. Parameter pada operator ini diatur “TF-IDF” pada opsi “vector creation” sesuai dengan pada gambar 11.



Gambar 11. Parameter Pada Operator “Process Document From Data”

Kemudian juga digunakan operator “Union” untuk menggabungkan data yang akan diuji dengan data latih yang telah dibuat. Setelah itu hubungkan kembali dengan operator “Filter Example” dengan kondisi “is missing” karena teks dengan sentimen kosong yang akan diuji dan diproses. Untuk mengisi kekosongan saat pemrosesan data diperlukan operator “Replace Missing Value” dengan kondisi default “zero” agar sistem tidak bingung dengan nilai data yang kosong. Setelah itu tambahkan operator “Apply Model” yang berfungsi untuk menyimpan dan menampilkan hasil sentimen yang dihasilkan dari pemrosesan Naïve Bayes.



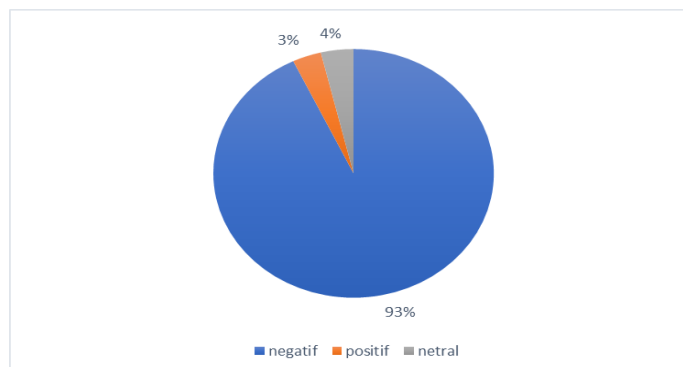
Gambar 12. Operator untuk Implementasi Naïve Bayes untuk Pengujian Data

Setelah dataset dimasukkan ke dalam RapidMiner untuk dilakukan proses cleansing dan preprocessing, terdapat frekuensi kata yang paling banyak muncul bentuk WordCloud. Pada penelitian ini, peneliti hanya mengambil sebanyak 20 kata dengan frekuensi kemunculan tertinggi. Semakin besar frekuensi munculnya kata, maka semakin besar juga ukuran kata dalam bentuk visualisasi. Seperti yang diperlihatkan pada gambar 13.



Gambar 13. Hasil WordCloud

Selanjutnya dataset yang sama diolah juga untuk mendapatkan analisis sentimen. Dari 905 data yang diolah, data dibagi menjadi dua bagian. Sebanyak 543 data dijadikan sebagai data latih, dan 362 dijadikan sebagai data uji. Pembagian jumlah data tersebut berpatokan pada perbandingan 60:40.



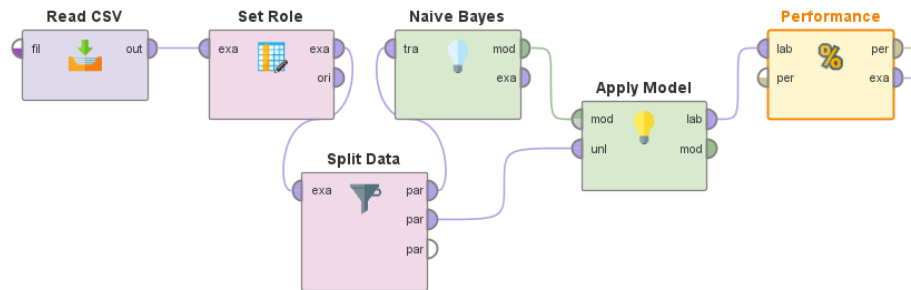
Gambar 14. Persentase Hasil Sentimen



Berdasarkan grafik persentase pada gambar 14 diatas, terdapat 905 data mengenai sentimen keberadaan juru parkir liar yang telah dianalisis. Didapatkan hasil sentimen negatif sebesar 93% atau sebanyak 841 sentimen, sentimen positif sebesar 3% atau sebanyak 30 sentimen. Dan sentimen netral sebesar 4% atau sebanyak 34 sentimen.

3.7 Evaluasi

Setelah melakukan penelitian berupa analisis sentimen pada dataset menggunakan metode Naïve Bayes, selanjutnya dilakukan pengujian Confusion Matrix. Pengujian ini bertujuan menghitung tingkat akurasi berdasarkan evaluasi dan validasi [28]. Confusion Matrix menghitung tingkat akurasi dengan tiga parameter, yakni accuracy, precision, dan recall [29]. Perhitungan akurasi ini menggunakan bantuan aplikasi Rapid Miner. Gambar 15 menunjukkan operator operator yang digunakan dalam proses perhitungan Confusion Matrix pada RapidMiner.



Gambar 15. Operator-operator yang digunakan untuk menghitung Confusion Matrix

Perhitungan Confusion Matrix diawali dengan menggunakan operator "Read CSV" yang berisi data latih kemudian dihubungkan dengan operator "Set Role". Operator "Set Role" berfungsi untuk mejadikan kolom hasil sentimen menjadi atribut label. Kemudian dihubungkan dengan operator "Split Data" yang berfungsi membagi data dengan pembagian 60:40 sesuai pada gambar 16.

ratio
0.6
0.4

Gambar 16. Pembagian Data Pada Operator "Split Data"

Pada parameter "Sampling type" menggunakan opsi "automatic" yang berarti data yang digunakan adalah acak. Kemudian dihubungkan kembali dengan operator "Naïve Bayes" dan "Apply Model" dan terakhir diakhiri dengan operator "Performance" untuk menyajikan hasil pengukuran akurasi. Hasil confusion matrix dengan nilai akurasi sebesar 92.82% seperti yang terlihat pada gambar 17.

accuracy: 92.82%				
	true negatif	true positif	true netral	class precision
pred. negatif	336	12	14	92.82%
pred. positif	0	0	0	0.00%
pred. netral	0	0	0	0.00%
class recall	100.00%	0.00%	0.00%	

Gambar 17. Hasil Pengujian Hasil dengan Confusion Matrix

Dengan perhitungan sebagai berikut:

a. $Accuracy = \frac{(Data\ Benar)}{(Jumlah\ Data)} = \frac{336}{362} = 0.92817$ atau 92.82% (1)

b. Precision

Untuk menghitung precision, digunakan tabel 5 dengan rumus sebagai berikut

$Precision = \frac{(TP)}{(TP + FP)}$ (2)

Tabel 5. Perhitungan Precision Masing Masing Kelas

	negatif	positif	netral
TP	336	0	0
FP	12 + 14	0	0
Precision	$336 / (336 + 26) = 0.92817$	0	0



c. Recall

Untuk menghitung Recall, digunakan tabel 6 dengan rumus sebagai berikut

$$\text{Recall} = \frac{(\text{TP})}{(\text{TP} + \text{FN})} \quad (3)$$

Tabel 6. Perhitungan Recall Masing Masing Kelas

	negatif	positif	netral
TP	336	0	0
FN	0 + 0	0	0
Recall	$336/(336+0) = 1$	0	0

4. KESIMPULAN

Setelah dilakukan penelitian menggunakan data dari internet berupa tweet warganet media sosial X dengan metode Naïve Bayes, dapat ditarik kesimpulan bahwasannya metode ini mendapatkan hasil akurasi yang tinggi yakni mencapai 92,82%. Adapun hasil dari 905 data yang diteliti, sebesar 93% atau sebanyak 841 sentimen menyatakan sentimen negatif, sebesar 3% atau sebanyak 30 sentimen yang menyatakan positif, dan sisanya sebesar 4% atau sebanyak 34 sentimen yang menyatakan netral dengan keberadaan juru parkir liar. Tingginya sentimen negatif dari warganet dikarenakan juru parkir liar kerap melakukan pungutan liar secara paksa dan sering kali tidak menjaga kendaraan di area parkir sebagaimana mestinya. Selain itu keberadaan juru parkir liar juga dipelopori oleh ormas setempat sehingga keberadaannya sulit sekali untuk ditertibkan oleh pihak yang berwenang.

UCAPAN TERIMA KASIH

Terima kasih secara penuh disampaikan kepada pihak-pihak yang telah bersedia membantu dan mendukung terlaksananya penelitian ini.

REFERENCES

- [1] T. Tranggono et al., "Peran Media Sosial Sebagai Wadah Aspirasi Masyarakat," *Bureaucracy Journal : Indonesia Journal of Law and Social-Political Governance*, vol. 3, no. 2, pp. 2155–2164, Jun. 2023, doi: 10.53363/BUREAU.V3I2.314.
- [2] Y. W. S. Putra et al., *Pengantar Aplikasi Mobile*. CV. Haura Utama, 2023.
- [3] Y. Uky et al., *Memahami Teknologi Informasi: Prinsip, Pengembangan, dan Penerapan*. Kaizen Media Publishing, 2023.
- [4] Alfandi Safira and F. N. Hasan, "Analisis Sentimen Masyarakat Terhadap Paylater Menggunakan Metode Naive Bayes Classifier," *ZONAsi: Jurnal Sistem Informasi*, vol. 5, no. 1, pp. 59–70, Jan. 2023, doi: 10.31849/zn.v5i1.12856.
- [5] Peraturan Republik Indonesia Nomor 112 Tahun 2007. Indonesia.
- [6] R. J. Oszaer et al., "Pencegakan Hukum Terhadap Juru Parkir Tidak Resmi Di Kota Ambon," *CAPITAN: Constitutional Law & Administrative Law Review*, vol. 1, no. 1, pp. 46–63, Jul. 2023, doi: 10.47268/CAPITAN.V1I1.9907.
- [7] S. E. Rahayu, H. Rita, and H. Febriaty, "Potensi Retribusi Parkir Terhadap Pendapatan Asli Daerah Kota Medan, Sebelum dan Sesudah Penerapan E-Parkir," *Owner : Riset dan Jurnal Akuntansi*, vol. 7, no. 4, pp. 3702–3711, Oct. 2023, doi: 10.33395/OWNER.V7I4.1936.
- [8] D. Noviantoro and A. F. Rosando, "Peran Dinas Perhubungan Dalam Menertibkan Oknum Jasa Parkir Liar Di Kota Surabaya," *Bureaucracy Journal : Indonesia Journal of Law and Social-Political Governance*, vol. 3, no. 2, pp. 1581–1591, Jan. 2023, doi: 10.53363/BUREAU.V3I2.266.
- [9] Kitab Undang-Undang Hukum Acara Pidana. Indonesia: Mahkamah Agung, 2021, p. 85.
- [10] Badan Pusat Statistik, "Februari 2023: Tingkat Pengangguran Terbuka (TPT) sebesar 5,45 Persen dan Rata-Rata Upah Buruh Sebesar 2,94 Juta Rupiah per Bulan." Accessed: Oct. 23, 2023. [Online]. Available: <https://www.bps.go.id/pressrelease/2023/05/05/2001/februari-2023--tingkat-pengangguran-terbuka--tpt--sebesar-5-45-persen-dan-rata-rata-upah-buruh-sebesar-2-94-juta-rupiah-per-bulan.html>
- [11] A. Hendri Doni, F. Alfiona, W. Andespa, E. dan Bisnis Islam, and U. M. Islam Negeri Sjech Djamil Djambek Bukittinggi, "Pengangguran Dalam Perspektif Ekonomi Islam Dan Kovensional," *Jurnal Ilmiah Ekonomi, Manajemen dan Syariah (JIEMAS)*, vol. 2, no. 1, pp. 1–10, Dec. 2022, doi: 10.55883/JIEMAS.V2I1.20.
- [12] R. Naquitasia, D. H. Fudholi, and L. Iswari, "Analisis Sentimen Berbasis Aspek pada Wisata Halal dengan Metode Deep Learning," *Jurnal Teknoinfo*, vol. 16, no. 2, pp. 156–164, Jul. 2022, doi: 10.33365/JTI.V16I2.1516.
- [13] C. F. Hasri and D. Alita, "Penerapan Metode Naïve Bayes Classifier Dan Support Vector Machine Pada Analisis Sentimen Terhadap Dampak Virus Corona Di Twitter," *Jurnal Informatika dan Rekayasa Perangkat Lunak (JATIKA)*, vol. 3, no. 2, pp. 145–160, Oct. 2022, doi: 10.33365/JATIKA.V3I2.2026.
- [14] E. Febriyani and H. Febriariyanti, "Analisis Sentimen Terhadap Program Kampus Merdeka Menggunakan Algoritma Naive Bayes Classifier Di Twitter," *Jurnal Tekno Kompak*, vol. 17, no. 1, pp. 25–38, Feb. 2023, doi: 10.33365/JTK.V17I1.2061.
- [15] T. A. Sari, E. Sinduningrum, and F. N. Hasan, "Analisis Sentimen Ulasan Pelanggan Pada Aplikasi Fore Coffee Menggunakan Metode Naïve Bayes," *KLIK: Kajian Ilmiah Informatika dan Komputer*, vol. 3, no. 6, pp. 773–779, Jun. 2023, doi: 10.30865/KLIK.V3I6.884.



- [16] N. Q. Rizkina and F. N. Hasan, "Analisis Sentimen Komentar Netizen Terhadap Pembubaran Konser NCT 127 Menggunakan Metode Naive Bayes," *Journal of Information System Research (JOSH)*, vol. 4, no. 4, pp. 1136–1144, Jul. 2023, doi: 10.47065/JOSH.V4I4.3803.
- [17] Y. Akbar and T. Sugiharto, "Analisis Sentimen Pengguna Twitter di Indonesia Terhadap ChatGPT Menggunakan Algoritma C4.5 dan Naive Bayes," *Jurnal Sains dan Teknologi*, pp. 115–122, 2023, doi: 10.55338/saintek.v4i3.1368.
- [18] S. Dwi et al., "Analisis Sentimen Relokasi Ibukota Nusantara Menggunakan Algoritma Naive Bayes dan KNN," *Jurnal KomtekInfo*, vol. 10, no. 1, pp. 1–7, Jan. 2023, doi: 10.35134/KOMTEKINFO.V10I1.330.
- [19] A. Saputra and F. N. Hasan, "Analisis Sentimen Terhadap Aplikasi Coffe Meets Bagel Dengan Algoritma Naive Bayes Classifier," *SIBATIK JOURNAL: Jurnal Ilmiah Bidang Sosial, Ekonomi, Budaya, Teknologi, Dan Pendidikan*, vol. 2, no. 2, pp. 465–474, Jan. 2023, doi: 10.54443/SIBATIK.V2I2.579.
- [20] T. Ridwansyah, "Implementasi Text Mining Terhadap Analisis Sentimen Masyarakat Dunia Di Twitter Terhadap Kota Medan Menggunakan K-Fold Cross Validation Dan Naive Bayes Classifier," *KLIK: Kajian Ilmiah Informatika dan Komputer*, vol. 2, no. 5, pp. 178–185, Apr. 2022, doi: 10.30865/KLIK.V2I5.362.
- [21] M. Hudha, E. Supriyati, and T. Listyorini, "Analisis Sentimen Pengguna Youtube Terhadap Tayangan #MATANAJWAMENANTITERAWAN Dengan Metode Naive Bayes Classifier," *JIKO (Jurnal Informatika dan Komputer)*, vol. 5, no. 1, pp. 1–6, Apr. 2022, doi: 10.33387/JIKO.V5I1.3376.
- [22] A. C. Khotimah and E. Utami, "Perbandingan Algoritma Naive Bayes Classifier, K-Nearest Neighbor Dan Support Vector Machine Dalam Klasifikasi Karakter Individu Pada Akun Twitter," *Jurnal Teknik Informatika (Jutif)*, vol. 3, no. 3, pp. 673–680, 2022, doi: 10.20884/1.jutif.2022.3.3.254.
- [23] A. R. Abdillah and F. N. Hasan, "Analisis Sentimen Terhadap Kandidat Calon Presiden Berdasarkan Tweets Di Sosial Media Menggunakan Naive Bayes Classifier," *SMATIKA JURNAL : STIKI Informatika Jurnal*, vol. 13, no. 01, pp. 117–130, Jul. 2023, doi: 10.32664/SMATIKA.V13I01.750.
- [24] A. R. Isnain, H. Sulistiani, B. M. Hurohman, A. Nurkholis, and Styawati, "Analisis Perbandingan Algoritma LSTM dan Naive Bayes untuk Analisis Sentimen," *JEPIN (Jurnal Edukasi dan Penelitian Informatika)*, vol. 8, no. 2, pp. 299–303, Aug. 2022, doi: 10.26418/JP.V8I2.54704.
- [25] A. Wibowo, F. Noor Hasan, L. Akbar Ramadhan, R. Nurhayati, and dan Arief Wibowo, "Analisis Sentimen Opini Masyarakat Terhadap Keefektifan Pembelajaran Daring Selama Pandemi COVID-19 Menggunakan Naive Bayes Classifier," *Jurnal Asimetri: Jurnal Ilmiah Rekayasa dan Inovasi*, vol. 4, pp. 239–248, Jul. 2022, doi: 10.35814/ASIIMETRIK.V4I1.3577.
- [26] D. R. Berliana and B. Santoso, "Elektabilitas Ridwan Kamil Dan Anies Baswedan Dalam Simulasi Pilpres 2024 Di Twitter (Analisis Jaringan Media Sosial Dan Analisis Sentimen Pengguna Twitter Terhadap #RIDWANKAMIL Dan #ANIESBASWEDAN)," *Mediakom : Jurnal Ilmu Komunikasi*, vol. 6, no. 2, pp. 150–162, Jan. 2023, doi: 10.35760/MKM.2022.V6I2.6962.
- [27] "Indonesian Stoplist." Accessed: May 23, 2023. [Online]. Available: <https://www.kaggle.com/datasets/oswinrh/indonesian-stoplist/data>
- [28] L. A. Rahman Malik, M. Kamayani, and F. N. Hasan, "Faktor-Faktor Yang Mempengaruhi Minat Calon Mahasiswa Baru Mendaftar Pada FTII Uhamka Menggunakan Algoritma K-Nearest Neighbor (K-NN)," *Infotech: Journal of Technology Information*, vol. 9, no. 1, pp. 85–94, Jun. 2023, doi: 10.37365/jti.v9i1.163.
- [29] B. Gunawan, H. Sasty, P. #2, E. Esyudha, and P. #3, "Sistem Analisis Sentimen pada Ulasan Produk Menggunakan Metode Naive Bayes," *JEPIN (Jurnal Edukasi dan Penelitian Informatika)*, vol. 4, no. 2, pp. 113–118, Dec. 2018, doi: 10.26418/JP.V4I2.27526.