

Implementasi Algoritma VADER terhadap Sentimen Larangan Mudik di Twitter

Mil Al Mizan
Universitas Buana Perjuangan
Karawang, Indonesia
ifl7.milmizan@mhs.ubpkarawang.ac.id

Hanny Hikmayanti H.
Universitas Buana Perjuangan
Karawang, Indonesia
hanny.hikmayanti@ubpkarawang.ac.id

Anis Fitri Nur Masruriyah
Universitas Buana Perjuangan
Karawang, Indonesia
anis.masruriyah@ubpkarawang.ac.id

Abstract— Kasus Covid-19 pertama di Indonesia diumumkan pada tanggal 2 Maret 2019. Pandemi Covid-19 memberikan dampak negatif terhadap berbagai bidang kehidupan seperti menurunnya aktivitas ekonomi masyarakat, meningkatnya angka kematian, dan memengaruhi sistem pendidikan. Berbagai upaya telah dilakukan pemerintah untuk menyelesaikan pandemi Covid-19, salah satunya dengan kebijakan larangan mudik. Dalam menanggapi kebijakan ini, masyarakat mempunyai tanggapannya masing-masing. Penelitian ini bertujuan untuk mengidentifikasi sentimen masyarakat terkait kebijakan larangan mudik Lebaran. Metode yang digunakan dalam penelitian ini adalah algoritma VADER. Data yang digunakan diambil dari media sosial Twitter sebanyak 1.000 cuitan, dikumpulkan pada periode waktu 6 Mei sampai 17 Mei 2021. Data terlebih dahulu melewati tahap preprocessing sebelum dinilai menggunakan algoritma VADER. Setelah itu, data diuji menggunakan confusion matrix. Hasil dari penelitian ini menunjukkan sentimen positif sebesar 42,2%, negatif 34,2%, dan netral 23,6%. Hasil dari pengujian data menunjukkan accuracy sebesar 50,3%, precision 54,81%, dan recall 58,02%.

Kata kunci — Analisis Sentimen, Confusion Matrix, VADER

I. PENDAHULUAN

Kasus Covid-19 pertama diumumkan di Indonesia pada tanggal 2 Maret 2019 [1]. Pandemi Covid-19 sulit diatasi, salah satu faktor yang menjadi penyebab yaitu kurangnya kepedulian masyarakat dalam menjaga protokol kesehatan [2]. Kesadaran diri sangat penting untuk dimiliki setiap orang dalam menjaga kesehatan, terutama di masa pandemi ini. Tindakan yang dilakukan untuk mencegah penyebaran Covid-19 adalah dengan cara mencuci tangan dengan sabun, memakai masker, dan menjaga jarak [3].

Berbagai upaya telah dilakukan pemerintah untuk menyelesaikan pandemi ini. Salah satu upaya pemerintah yaitu dengan menerapkan kebijakan larangan mudik. Larangan mudik 2021 diberlakukan mulai tanggal 6 Mei sampai 17 Mei 2021, hal ini tertulis dalam Surat Edaran Nomor 13 Tahun 2021 [4]. Dalam menanggapi kebijakan ini, masyarakat mempunyai opininya masing-masing; ada yang menanggapi secara netral, positif, dan ada juga yang negatif. Maka dari itu, perlu dilakukan analisis sentimen masyarakat terkait kebijakan larangan mudik 2021.

Analisis sentimen merupakan metode pengolahan data secara otomatis untuk mendapatkan informasi terkait sentimen positif, negatif, atau netral [5]. Penelitian terkait analisis sentimen telah banyak dilakukan, salah satunya yaitu kebijakan terkait larangan mudik pada tahun 2020. Data analisis sentimen diambil dari media sosial Twitter dengan metode Naive Bayes, hasilnya yaitu positif 199 cuitan dan negatif 34 cuitan dengan tingkat akurasi 83,83% [6]. Penelitian tentang sentimen pengadaan vaksin Covid-19 menunjukkan bahwa hasil penelitian ini adalah sentimen positif 48% (471 cuitan), negatif 23% (224 cuitan), dan netral 29% (281 cuitan), yang dapat disimpulkan bahwa hasilnya positif [7]. Kemudian, pada penelitian mengenai gerakan anti-vaksin, data diambil dari Twitter kemudian dikelompokkan menggunakan algoritma K-Means, lalu dianalisis dengan algoritma VADER (Valence Aware Dictionary for sEntiment Reasoning). Hasil yang didapatkan adalah pengklasteran data menjadi 10 klaster, setiap klaster diberi sentimen berdasarkan kata kunci yang sering muncul [8].

Dalam analisis sentimen yang akan dilakukan, data untuk penelitian ini akan diambil dari media sosial Twitter. Pada bulan Januari 2021, Indonesia merupakan peringkat ke-7 dunia dalam jumlah pengguna Twitter dengan jumlah pengguna mencapai 14,05 juta pengguna [9]. Twitter sering digunakan masyarakat dalam menyampaikan opini tentang berbagai macam hal. Hal ini dapat dimanfaatkan, misalnya bagi pemerintah dalam mengevaluasi kebijakan yang akan, sedang, atau telah diterapkan.

II. DATA DAN METODE

A. Bahan Penelitian

Bahan untuk penelitian ini diperoleh dari cuitan berbahasa Indonesia yang terdapat di media sosial Twitter dengan kata kunci “larangan mudik” menggunakan library Snsrape Python. Data ini diambil dari tanggal 6 Mei sampai 17 Mei 2021, berkaitan dengan larangan mudik 2021, yang dibatasi sebanyak 1000 cuitan. Kemudian, data diberi label sentimen dengan menggunakan algoritma VADER.

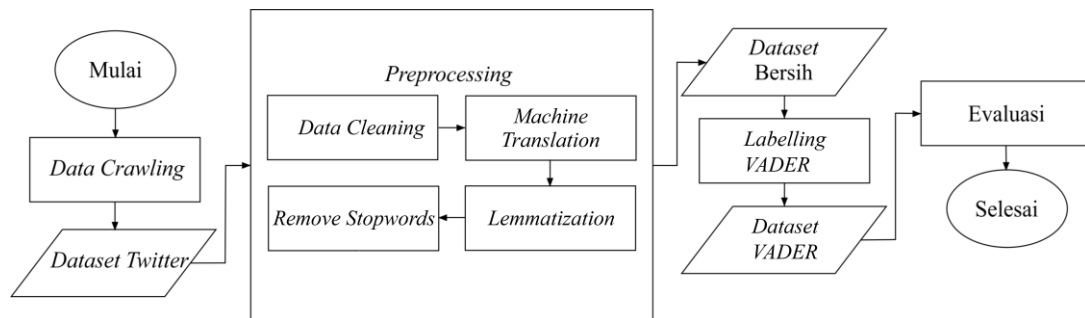
B. Prosedur Penelitian

Prosedur penelitian mempunyai tahapan proses seperti pada Gambar 3.1 yang dimulai dengan data crawling, yaitu tahap pengambilan data di Twitter. Kemudian, dataset diolah melalui tahap preprocessing yang terdiri dari data cleaning (pembersihan data dari simbol-simbol yang tidak diperlukan) dan penyeragaman bentuk huruf menjadi huruf kecil semua agar data mudah dibaca.

Lalu machine translation, yaitu data diterjemahkan dari bahasa Indonesia ke bahasa Inggris. Selanjutnya lemmatization, yaitu penyederhanaan bentuk kata menjadi bentuk standar atau asli dengan menggunakan kamus. Remove stopwords, yaitu proses menghapus kata tambahan atau kata penghubung.

Setelah dataset melewati tahap preprocessing, kemudian dataset melalui tahap labelling VADER, yaitu proses pemberian label sentimen berupa negatif, netral, atau positif. Kemudian dataset yang telah diberi label dievaluasi dengan menggunakan confusion matrix.

Prosedur pada penelitian ini adalah sebagai berikut:



Gambar 1 Prosedur Penelitian

III. HASIL DAN PEMBAHASAN

A. Data Crawling

Hasil dari data crawling dibatasi sebanyak 1000 data, diambil dari periode waktu 6 Mei sampai 17 Mei 2021 dengan kata kunci “larangan mudik”. Data ini masih berupa data mentah yang akan melewati beberapa tahapan untuk menghilangkan data yang tidak relevan.

Tabel 1 Data Crawling

No.	Content
1	Kemarin 36.468 Kendaraan Diputarbalik Terkait Larangan Mudik https://t.co/tyTHvz6wd2
2	Saat ini marak pelanggar prokes/aturan larangan mudik videonya diviralkan. Sampai identitas pribadinya disebar kemana2. Pelaku dibully habis2an oleh netizen. Dibuat malu sampai akhirnya minta maaf di depan media dan viral juga. Apakah dgn cara2 ini orang jadi makin taat 3M/5M?
3	Pemudik disuruh putar balik ngomel-ngomel, ngamuk, maki2 petugas. Ya disuruh putar balik itu resiko dari mudik di tengah larangan. Gimana sih warga?! Tahan, tahan bentar. Corona masih ada #Lebaran2021 #Lebaran2021TidakMudik
4	Sobat Jatim, yuk kita simak bersama DINAMIKA JAWA TIMUR Senin, 17 Mei 2021 Pukul 07.00 – 08.00 WIB Dengan tema : "MENGUKUR EFEKTIFITAS PEMBATAHAN DAN LARANGAN MUDIK". @KhofifahIP @EmilDardak @KominfoJatim https://t.co/h5DyLhDmly
5	@DioR3born Patuhilah larangan mudik
...	...
996	Polisi: 1 Juta Orang Masuk-Keluar Jabodetabek Sebelum Larangan Mudik Diberlakukan https://t.co/4fhJ65iQTH
997	@deYOUUniverse Selamat malam, sesuai surat dari Bupati Lebak tanggal 29 April 2021 nomor 443/1684-GT/2021, maka pada masa larangan mudik lebaran ini Stasiun, Maja, Citeras, dan Rangkasbitung tidak melayani naik-turun pengguna KRL 1
998	berasa juga imbas dari larangan mudik, bapak w paling tua di antara keluarga yang ada di jakarta dan semuanya gak bisa mudik jadi ngumpul di rumah w ga udah-udah. dari yang taunya 'ende' w cuma 2 ini jadi 13 org???!!!!
	dan besok masih ada cloter selanjutnya, mantap.
999	@lazy lucky99 Sabar ya kak, larangan mudik masih berlaku soalnya sampai tanggal 17 Mei, bagi yang nekat mudik akan diberikan sanksi berupa putar balik, denda dan dipidana bagi mobil pribadi yang jadi travel gelap.
1000	@ryan_nus Wahai CEO @gojekindonesia .

anda noob sekali sampai gak bisa antisipasi lonjakan di libur lebaran efek larangan mudik. cuman nambahin insentif sedikit buat seminggu ini aja gak bisa di estimasi, itung2 bikin semangat driver.

B. Preprocessing

1. Data Cleaning

Data cleaning digunakan untuk menyeragamkan bentuk huruf menjadi huruf kecil dan menghilangkan data yang tidak diperlukan seperti angka, username, simbol, dan sebagainya.

Tabel 2 Data Cleaning

No.	Data Cleaning
1	kemarin kendaraan diputarbalik terkait larangan mudik
2	saat ini marak pelanggaran prokes aturan larangan mudik videonya diviralkan sampai identitas pribadinya disebar kemana pelaku dibully habis an oleh netizen dibuat malu sampai akhirnya minta maaf di depan media dan viral juga apakah dgn cara ini orang jadi makin taat m m
3	pemudik disuruh putar balik ngomel ngomel ngamuk maki petugas ya disuruh putar balik itu resiko dari mudik di tengah larangan gimana sih warga tahan tahan bentar corona masih ada lebaran lebaran tidakmudik
4	sobat jatim yuk kita simak bersama dinamika jawa timur senin mei pukul wib dengan tema mengukur efektifitas pembatasan dan larangan mudik
5	patuhilah larangan mudik
...	...
996	polisi juta orang masuk keluar jabodetabek sebelum larangan mudik diberlakukan
997	selamat malam sesuai surat dari bupati lebak tanggal april nomor gt maka pada masa larangan mudik lebaran ini stasiun maja citeras dan rangkasbitung tidak melayani naik turun pengguna krl
998	berasa juga imbas dari larangan mudik bapak w paling tua di antara keluarga yang ada di jakarta dan semuanya gak bisa mudik jadi ngumpul di rumah w ga udah udah dari yang taunya ende w cuma ini jadi org dan besok masih ada cloter selanjutnya mantap
999	sabar ya kak larangan mudik masih berlaku soalnya sampai tanggal mei bagi yang nekat mudik akan diberikan sanksi berupa putar balik denda dan dipidana bagi mobil pribadi yang jadi travel gelap
1000	wahai ceo anda noob sekali sampai gak bisa antisipasi lonjakan di libur lebaran efek larangan mudik cuman nambahin insentif sedikit buat seminggu ini aja gak bisa di estimasi itung bikin semangat driver

2. Machine Translation

Machine translation diperlukan karena algoritma VADER menggunakan kamus leksikon berbahasa Inggris. Selain itu, proses ini digunakan agar proses pelabelan sentimen menjadi lebih cepat dan mudah. Proses ini menggunakan library googletrans.

Tabel 3 Machine Translation

No.	Machine Translation
1	yesterday the vehicle was turned around related to the restriction of homecoming
2	Currently the rampant violators of the processed rule for the video homecoming of the video are visited until his personal identity is spread where the perpetrators were bullied by netizens made embarrassed until finally apologizing in front of the media and viral as well whether in this way people became more obedient to
3	The travelers were told to turn back and grumble nagging the officers of the officers who were told to turn back the risk of going home in the midst of the prohibition how the residents resistant to the time of Corona still there was Eid Lebaran
4	Let's friends, let's see with the dynamics of East Java Monday May at WIB with the theme of measuring the effectiveness of restrictions and restrictions on homecoming
5	Adhere to mudic prohibition
...	...
996	Police Million People Enter Jabodetabek before the home to the homecoming was imposed
997	Good evening according to the letter from the Regent of Lebak on April Number GT, then during the Lebaran homecoming period the Maja Citeras Station and Rangkasbitung did not serve up and down KRL users
998	It also feels the impact of the rest of the homecoming of Mr. W the oldest among the families in Jakarta and all of them can't go home so they gather at home w, it's already from the ende w.
999	Patience, sis, the restrictions on the homecoming are still valid because until the May date for those who are determined to go home, they will be given sanctions in the form of a fine and priced for a private car that becomes a dark travel
1000	O your CEO, noob so that you can't anticipate the surge in the Lebaran holiday, the effects of going home just adding a little incentive for this week.

3. Lemmatization

Pada tahapan lemmatization dilakukan proses penyederhanaan bentuk kata menjadi kata dasar berdasarkan kamus yang ada di library Python NLTK. Sebenarnya tahap ini sudah dilakukan saat di tahap machine translation, namun tetap dilakukan untuk memastikan tidak terjadi kesalahan data saat proses penerjemahan.

Tabel 4 Lemmatization

No.	Lemmatization
1	yesterday the vehicle was turned around related to the restriction of homecoming
2	Currently the rampant violators of the processed rule for the video homecoming of the video are visited until his personal identity is spread where the perpetrators were bullied by netizens made embarrassed until finally apologizing in front of the media and viral as well whether in this way people became more obedient to
3	The travelers were told to turn back and grumble nagging the officers of the officers who were told to turn back the risk of going home in the midst of the prohibition how the residents resistant to the time of Corona still there was Eid Lebaran
4	Let's friends, let's see with the dynamics of East Java Monday May at WIB with the theme of measuring the effectiveness of restrictions and restrictions on homecoming
5	Adhere to mudic prohibition
...	...
996	Police Million People Enter Jabodetabek before the home to the homecoming was imposed
997	Good evening according to the letter from the Regent of Lebak on April Number GT, then during the Lebaran homecoming period the Maja Citeras Station and Rangkasbitung did not serve up and down KRL users
998	It also feels the impact of the rest of the homecoming of Mr. W the oldest among the families in Jakarta and all of them can't go home so they gather at home w, it's already from the end w.
999	Patience, sis, the restrictions on the homecoming are still valid because until the May date for those who are determined to go home, they will be given sanctions in the form of a fine and priced for a private car that becomes a dark travel
1000	O your CEO, noob so that you can't anticipate the surge in the Lebaran holiday, the effects of going home just adding a little incentive for this week.

4. Remove Stopwords

Tahapan untuk menghilangkan kata imbuhan yang tidak diperlukan dilakukan dengan menggunakan library Python NLTK.

Tabel 5 Remove Stopwords

No.	Remove Stopwords
1	yesterday vehicle turned around related restriction homecoming
2	Currently rampant violators processed rule video homecoming video visited personal identity spread perpetrators bullied netizens made embarrassed finally apologizing front media viral well whether way people became obedient
3	The travelers told turn back grumble nagging officers officers told turn back risk going home midst prohibition residents resistant time Corona still Eid Lebaran
4	Let's friends, let's see dynamics East Java Monday May WIB theme measuring effectiveness restrictions restrictions homecoming
5	Adhere mudic prohibition
...	...
996	Police Million People Enter Jabodetabek home homecoming imposed
997	Good evening according letter Regent Lebak April Number GT, Lebaran homecoming period Maja Citeras Station Rangkasbitung serve KRL users
998	It also feels impact rest homecoming Mr. W oldest among families Jakarta can't go home gather home w, already end w.
999	Patience, sis, restrictions homecoming still valid May date determined go home, given sanctions form fine priced private car becomes dark travel
1000	O CEO, noob can't anticipate surge Lebaran holiday, effects going home adding little incentive week.

C. Labelling VADER

Labelling VADER dibagi menjadi 3 kategori: netral, positif, dan negatif. Skor 0 masuk kategori netral, skor lebih dari 0,05 masuk dalam kategori positif, dan skor kurang dari -0,05 masuk dalam kategori negatif. Dari 1000 data yang digunakan pada penelitian ini, didapatkan hasil prediksi sentimen: positif 422, negatif 342, dan netral 236.

Tabel 6 Labelling VADER

No.	Content	Score	Sentiment
1	yesterday vehicle turned around related restriction homecoming	-0.2732	Negative
2	Currently rampant violators processed rule video homecoming video visited personal identity spread perpetrators bullied netizens made embarrassed finally apologizing front media viral well whether way people became obedient	-0.8658	Negative
3	The travelers told turn back grumble nagging officers officers told turn back risk going home midst prohibition residents resistant time Corona still Eid Lebaran	-0.5859	Negative
4	Let's friends, let's see dynamics East Java Monday May WIB theme measuring effectiveness restrictions restrictions homecoming	0.6369	Positive
5	Adhere mudic prohibition	0.0	Neutral
...	...	...	...
996	Police Million People Enter Jabodetabek home homecoming imposed	-0.0772	Negative
997	Good evening according letter Regent Lebak April Number GT, Lebaran homecoming period Maja Citeras Station Rangkasbitung serve KRL users	0.7213	Positive
998	It also feels impact rest homecoming Mr. W oldest among families Jakarta can't go home gather home w, already ende w.	0.0	Neutral
999	Patience, sis, restrictions homecoming still valid May date determined go home, given sanctions form fine priced private car becomes dark travel	0.4939	Positive
1000	O CEO, noob can't anticipate surge Lebaran holiday, effects going home adding little incentive week.	0.5729	Positive

D. Pengujian

Pada proses pengujian, penelitian ini menggunakan metode confusion matrix. Bagian sisi predicted digunakan untuk menampilkan hasil prediksi yang dilakukan oleh algoritma VADER, sedangkan pada sisi actual digunakan untuk menampilkan hasil sentimen yang asli berdasarkan validasi dari ahli bahasa Indonesia.

Tabel 4 *Confusion Matrix*

	<i>Actual Positive</i>	<i>Actual Negative</i>	<i>Actual Neutral</i>
<i>Predicted Positive</i>	106	142	174
<i>Predicted Negative</i>	15	220	107
<i>Predicted Neutral</i>	8	51	177

$$1. \text{ Accuracy} = \frac{106+220+177}{106+15+8+142+220+51+174+107+177} = 0.503 \text{ atau jika dalam persentase } 50.3\%$$

2. Precision

$$\text{Positive} = \frac{106}{106+142+174} = 0.2511$$

$$\text{Negative} = \frac{220}{220+15+107} = 0.6432$$

$$\text{Neutral} = \frac{177}{177+8+51} = 0.75$$

$$\text{Rata-rata nilai precision} = (0.2511+0.6432+0.75)/3 = 0.5481 \text{ atau jika dalam persentase } 54.81\%$$

3. Recall

$$\text{Positive} = \frac{106}{106+15+8} = 0.8217$$

$$\text{Negative} = \frac{220}{220+142+51} = 0.5326$$

$$\text{Neutral} = \frac{177}{177+107+174} = 0.3864$$

$$\text{Rata-rata nilai recall} = (0.8217+0.5326+0.3864)/3 = 0.5802 \text{ atau jika dalam persentase } 58.02\%$$

IV. KESIMPULAN

Hasil analisis sentimen masyarakat terhadap larangan mudik 2021 dengan menggunakan algoritma VADER menunjukkan sentimen positif sebesar 42,2%, negatif 34,2%, dan netral 23,6%. Hasil pengujian analisis sentimen menggunakan confusion matrix menghasilkan akurasi sebesar 50,3%, presisi 54,81%, dan recall 58,02%. Hasil yang tergolong rendah ini disebabkan oleh beberapa kata yang sebenarnya bersifat negatif namun diklasifikasikan sebagai positif karena perbedaan penilaian sentimen yang bergantung pada konteks. Selain itu, klasifikasi untuk sentimen netral memang lebih sulit dibandingkan dengan sentimen negatif dan positif.

1. PENGAKUAN

Naskah ilmiah ini merupakan sebagian dari penelitian pada tugas akhir milik Mil Al Mizan dengan judul Analisis Sentimen Larangan Mudik pada Media Sosial Twitter, yang dibimbing oleh Dr. Hanny Hikmayanti H., M.Kom. dan Anis Fitri Nur Masruriyah, M.Kom.

2. DAFTAR PUSTAKA

- [1] CNN, “Jokowi Umumkan Dua WNI Positif Corona di Indonesia,” 2020. <https://www.cnnindonesia.com/nasional/20200302111534-20-479660/jokowi-umumkan-dua-wni-positif-corona-di-indonesia> (accessed Jun. 15, 2021).
- [2] R. Fadillah Aulia, Elvina, Farhan, Fajri, Lisa, Neo, Nurun, “Kekalahan Indonesia Dalam Perang Pandemi Covid-19,” *Academia*, 2020, [Online]. Available: https://www.academia.edu/43806583/KEKALAHAN_INDONESIA_DALAM_PERANG_PANDEMI_COVID_19
- [3] Kementerian Kesehatan, “Pedoman Pencegahan dan Pengendalian Corona Virus deases (Covid-19),” *Kementrian Kesehatan*, vol. 5. p. 178, 2020. [Online]. Available: https://covid19.go.id/storage/app/media/Protokol/REV-05_Pedoman_P2_COVID-19_13_Juli_2020.pdf
- [4] Satuan Tugas Penanganan COVID-19, “Surat Edaran Ka Satgas Nomor 13 Tahun 2021 Larangan Mudik Hari Raya Idul Fitri dan Pengendalian COVID-19 selama Bulan Suci Ramadhan 1442H.” p. 7, 2021.
- [5] F. V. Sari and A. Wibowo, “Analisis Sentimen Pelanggan Toko Online Jd.Id Menggunakan Metode Naïve Bayes Classifier Berbasis Konversi Ikon Emosi,” *J. SIMETRIS*, vol. 10, no. 2, pp. 681–686, 2019.
- [6] T. A. M, Y. Alkhalifi, N. A. Mayangky, and W. Gata, “Analisis Sentimen Opini Publik Mengenai Larangan Mudik pada Twitter Menggunakan Naive Bayes,” *CoreIT*, vol. 6, no. 2. 2020.
- [7] N. Rakhmawati, M. Aditama, R. Pratama, and K. Wiwaha, “Analisis Klasifikasi Sentimen Pengguna Media Sosial Twitter Terhadap Pengadaan Vaksin COVID-19,” *JIEET (Journal Inf. Eng. Educ. Technol.)*, vol. 4, no. 2, pp. 90–92, 2020.
- [8] J. Garay, R. Yap, and M. J. Sabellano, “An analysis on the insights of the anti-vaccine movement from social media posts using k-means clustering algorithm and VADER sentiment analyzer,” *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 482, no. 1, pp. 0–6, 2019, doi: 10.1088/1757-899X/482/1/012043.
- [9] H. Tankovska, “Leading countries based on number of Twitter users as of January 2021,” 2021. [Online]. Available: <https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/>