

ANALISIS SENTIMEN MASYARAKAT PADA MEDIA SOSIAL TWITTER TERHADAP PENERAPAN NEW NORMAL DI INDONESIA MENGGUNAKAN METODE LEXICON BASED

Arnoldus Janssen Dahur¹, Isa Albanna²

Intitut Teknologi Adhi Tama Surabaya

arnolddahur@gmail.com, isaalbanna@itats.ac.id

ABSTRACT

Covid-19 Pandemic has brought great impacts to the development of human life almost for all countries in the world including Indonesia. Every day, there are people who get diseases and finally die caused by Covid-19. Therefore, activities in various aspects could not run normally as usual. To overcome this sort of condition, many countries have issued the policy of new normal. Unfortunately, people convey pros and cons against this policy by stating their opinions in the social media such as twitter. Basically, these opinions can be used as information to formulate the next government policy. To process the data of people's opinions revealed through twitter social media, sentiment analysis is necessary to conduct. Several stages in sentiment analysis involved data collection, pre-processing, classifying, and visualizing. For this reason, the researcher employed a lexicon-based method in the process of data classification. The result of final classification demonstrated that out of 1338 total tweets, 563 data were positive, 751 negative, and 24 neutral. In the other words, when the data were converted into percentages, they showed 42.1 % positive sentiment, 56.1% negative sentiment, and 1.8% neutral sentiment.

Keywords: New Normal, Classifier, Lexicon Based, prediction, Twitter

ABSTRAK – Font 10

Pandemi virus covid-19 memberikan banyak dampak bagi perkembangan kehidupan manusia disemua bagian negara di dunia., tak luput juga di Indonesia. Setiap harinya ada korban manusia yang terjangkit virus covid-19 yang menyebabkan banyak korban meninggal dunia. Oleh karena itu menyebabkan aktivitas dalam berbagai aspek banyak yang tidak berjalan normal seperti biasanya, untuk mengatasi itu semua banyak negara mengeluarkan kebijakan masing-masing dan yang banyak digunakan adalah kebijakan new normal. Akan tetapi kebijakan new normal ini banyak menuai pro dan kontra di tengah masyarakat yang menyebabkan berbagai opini dari mereka yang seringkali mereka tuangkan lewat media social, salah satunya media sosial twitter. Oleh karena itu opini masyarakat itu bisa menjadi informasi yang digunakan untuk mengeluarkan kebijakan pemerintah selanjutnya. Untuk mengolah data dari opini masyarakat yang mereka tuangkan lewat media sosial twitter maka dilakukan analisis sentiment. Analisis sentiment ini memiliki tahapan-tahapan didalamnya diantaranya pengumpulan data, preprocessing, klasifikasi, dan yang terakhir visualisasi data. Dalam penelitian ini digunakan metode lexicon based untuk proses klasifikasi data. Berdasarkan penelitian tersebut maka didapat hasil akhir klasifikasi yaitu dari jumlah 1338 tweet total keseluruhan data, setelah dianalisis ada 563 data yang positif, 751 yang negative dan ada 24 yang netral, atau dikonversikan dalam bentuk persen, yaitu 42,1 % untuk sentiment positif, 56,1% untuk sentiment negative, dan 1,8% untuk sentiment netral.

Kata kunci: New Normal, Classifier, Lexicon Based, prediksi, Twitter.

PENDAHULUAN

Pandemi covid-19 telah memberikan dampak yang begitu luar biasa di seluruh dunia dan tidak luput juga bagi negara Indonesia. Terdapat begitu banyak korban jiwa yang disebabkan oleh pandemic covid-19. Menurut *World Health Organization (WHO)*, ada sekitar 30.000.000 populasi manusia yang terkena dampak covid-19 ini dan ada lebih dari 943.000 orang telah meninggal di seluruh dunia (*WHO | Organisasi Kesehatan Dunia, 2020/09/2020*). Di Indonesia sendiri menurut informasi satuan tugas (satgas) covid-19 sudah lebih dari 232.000 orang telah

tertular Covid-19 dan lebih dari 9200 orang telah meninggal dunia Banyaknya yang tertular dan korban jiwa yang disebabkan dari pandemic covid-19 ini, membuat hampir semua aktifitas sehari-hari tidak berjalan seperti biasanya. Banyak kegiatan-kegiatan yang harus dilakukan dirumah seperti kegiatan belajar mengajar, perkantoran dan masih banyak lagi lainnya. Oleh karena itu pemerintah harus menemukan solusi agar roda kehidupan tetap dapat berjalan ditengah pandemic ini, maka muncullah sebuah tatanan, kebiasaan dan perilaku yang baru berbasis pada adaptasi untuk membudayakan perilaku hidup bersih dan sehat inilah yang kemudian disebut sebagai "new normal"[1].

Mengacu pada paparan diatas agar data-data ini bisa dianalisa dibutuhkan sebuah metode yang dapat menentukan hasil akhir dari analisa ini. Metode-metode yang digunakan juga ada banyak, dan yang digunakan pada penelitian ini adalah metode *lexicon based*. Metode *lexicon based* merupakan salah satu metode berbasis kamus untuk mendukung klasifikasi sentiment dengan dibantu *SentiWordNet*. Dalam penelitian ini Kamus atau *lexicon* yang digunakan yaitu kamus Bahasa inggris. *SentiwordNet* itu sendiri yaitu basis data leksikal yang dibuat berguna dalam membantu pengklasifikasian sentiment dan proses-proses opinion mining [2].

TINJAUAN PUSTAKA

Analisis Sentiment

Analisis sentimen yaitu sebuah cabang keilmuan yang bermanfaat untuk menganalisis opini, sentimen, evaluasi, penilaian, sikap, dan emosi terhadap suatu entitas seperti produk, layanan, organisasi, individu, isu, peristiwa, dan topik [3]. Dijelaskan lagi dalam jurnalnya bahwa analisis sentiment merupakan salah satu cara untuk bisa menganalisis opini, sentiment, penilaian, perilaku, dan perasaan individu atau kelompok terhadap suatu kejadian, lembaga, perorangan.

Text Mining

Text mining yaitu cara menganalisa suatu dokumen, selanjutnya dikelompokkan berdasarkan kata-kata yang terkandung didalamnya, dan menentukan kesamaan satu sama lain untuk mengetahui keterkaitan antara kata-kata tersebut[4].

Preprocessing

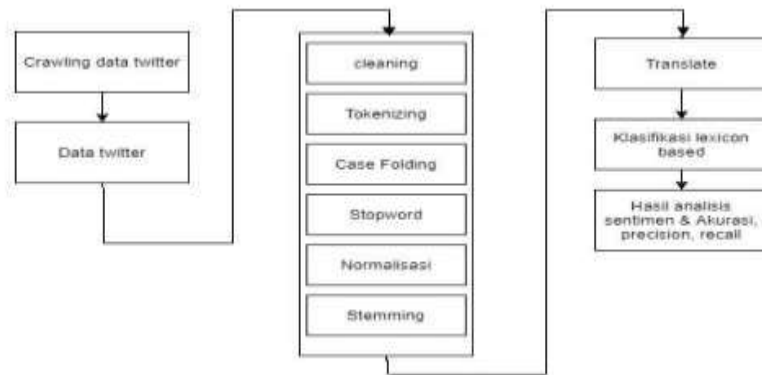
Preprocessing yaitu mengubah data atau dokumen teks yang belum terstruktur menjadi dokumen teks terstruktur yang siap digunakan untuk proses selanjutnya[5]. Adapun beberapa tahapan didalam preprocessing itu sendiri yaitu cleaning, case folding, tokenizing, stopword, normalisasi dan stemming

Lexicon Based

Lexicon based features ialah fitur atau kata yang telah diberi bobot berdasarkan kamus atau *lexicon*. Pemberian bobot dilakukan untuk setiap kata yang termasuk sentimen positif atau termasuk sentimen negatif . Pendekatan menggunakan kata opini untuk menentukan sentimen disebut pendekatan berbasis kamus (*lexicon-based aproach*)[6].

METODE

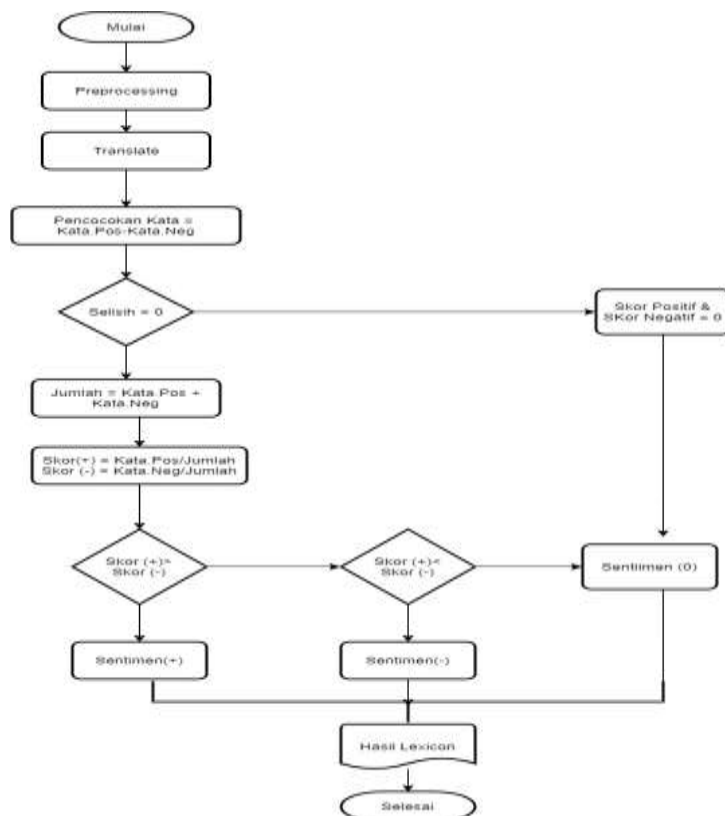
Tahapan proses metode penelitian dengan metode lexicon based



Gambar 1. Diagram Flowcart Tahapan Metode Penelitian

Tahapan preprocessing bertujuan untuk mengubah data twitter yang sudah dicrawling sebelumnya menjadi data yang lebih terstruktur, karena pada dasarnya data yang di crawling masih merupakan data tidak terstruktur. Tahapan preprocessing ini terbagi menjadi beberapa tahapan yaitu cleaning, tokenisasi, case folding, stopword, normalisasi bahasa, stemming

Lexicon Based



Gambar 2. Flowchart Klasifikasi Lexicon Based

Tabel 1. Contoh hasil klasifikasi Lexicon Based

No	Stemming	Sentimen Score	Hasil
1	['masuk', 'era', 'adaptasi', 'biasa', 'new', 'normal', 'wajib','ikut','protokol', 'sehat', 'jaga', 'jarak','masker']	<i>negative score</i> = - 0.375 <i>positivity score</i> = 0.5	kelas Positif
2	['swab', 'test', 'pilkada', 'serentak', 'moga', 'sadar', 'paslon', 'pilkada', 'era', 'new','normal']	<i>negative score</i> = - 0.175 <i>positivity score</i> = 0.6235	kelas positif
3	['era', 'new', 'normal', 'siap', 'lengkap', 'pribadi', 'sesuai', 'protokol', 'sehat', 'hindar', 'covid']	<i>negative score</i> = - 0.125 <i>positivity score</i> = 0.125	kelas netral

Jadi suatu kalimat akan dikatakan positif, negative dan netral dilihat dari :

$$Sentencesentiment \begin{cases} positive & \text{if } S_{positive} > S_{negative} \\ neutral & \text{if } S_{positive} = S_{negative} \\ negative & \text{if } S_{positive} < S_{negative} \end{cases}$$

HASIL DAN PEMBAHASAN

Tahapan-tahapan dalam penelitian ini diawali dengan pengumpulan data twitter dengan cara crawling data twitter dengan bantuan twitter API dan menggunakan phyton sebagai bahasa pemrograman dengan library jupyter notebook. Tahapan selanjutnya adalah preprocessing yang dimana bertujuan untuk mengubah data yang tidak terstruktur dari crawling data sebelumnya menjadi data yang lebih terstruktur. Langkah-langkah preprocessing sebagai berikut

Tabel 2. Contoh tahapan preprocessing

Komentar	Cleaning	Tokenizing	Case Folding	Normalisasi	Stopword	Stemming
Beradaptasi dgn new normal bisa memberikan rasa aman	Beradaptasi dgn new normal bisa memberikan rasa aman	Beradaptasi', 'dgn', 'new', 'normal', 'bisa', 'memberikan', 'rasa', 'aman'	beradaptasi', 'dgn', 'new', 'normal', 'bisa', 'memberikan', 'rasa', 'aman'	beradaptasi', 'dengan', 'new', 'normal', 'bisa', 'memberikan', 'rasa', 'aman'	beradaptasi', , 'new', 'normal', 'al', 'memberikan', 'rasa', 'aman'	adaptasi', 'new', 'normal', 'beri', 'rasa', 'aman'
Gagal terapkan new normal	Gagal terapkan new normal	Gagal', 'terapkan', 'new', 'normal'	gagal', 'terapkan', 'new', 'normal'	gagal', 'terapkan', 'new', 'normal'	gagal', 'terapkan', 'new', 'normal'	agal', 'terap', 'new', 'normal'

Contoh perhitungan menggunakan metode lexicon based

- $sum_kata = \sum_{i=1}^n PosScore(kata)$

$$sum_kata = \sum_{i=1}^n NegScore(kata)$$

$$sum_PosScore = 0,375 + 0,0 + 0,0 + 0,0 + 0,0 + 0,0 + 0,0 + 0,125 + 0,0 + 0,0 + 0,0 = 0,5$$

$$sum_NegScore = 0,0 + 0,5 + 0,0 + 0,125 + 0,0 + 0,0 + 0,75 + 0,0 + 0,0 + 0,0 + 0,0 = 1,375$$

Setelah dilakukan perhitungan seperti diatas maka didapat hasil klasifikasi secara keseluruhan yaitu

Tabel 3. Hasil Klasifikasi metode lexicon based

<i>Kategori</i>	<i>Jumlah</i>	<i>Presentase</i>
<i>Positif</i>	563	42,10%
<i>Netral</i>	24	56,10%
<i>Negatif</i>	751	1,80%

Berikutnya adalah melakukan perhitungan tingkat error dari hasil klasifikasi diatas dan akan mendapatkan hasil berupa tabel berikut :

Tabel 3. Tabel Confusion Matrix

<i>LEXICON</i>	<i>PREDIKSI</i>			
	<i>POSITIF</i>	<i>NEGATIF</i>	<i>NETRAL</i>	<i>TOTAL</i>
<i>POSITIF</i>	267	187	109	563
<i>NEGATIF</i>	178	324	249	751
<i>NETRAL</i>	6	4	14	24
<i>TOTAL</i>	451	515	372	

$$Accuracy = \frac{\sum pb}{num_data} \times 100\%$$

$$Accuracy = \frac{(total\ data - data\ error)}{total\ data} \times 100\%$$

$$Accuracy = \frac{(1338 - 659)}{1338} \times 100\%$$

$$Accuracy = 50,74\%$$

Tabel 4. Hasil Precision dan Recall

	Positif	Negatif	Netral
<i>Precision</i>	59,20177	62,91262	3,763441
<i>Recall</i>	47,42451	43,14248	58,33333

KESIMPULAN

Berdasarkan penelitian yang sudah dilakukan adapun kesimpulan yang diberikan dari penelitian yaitu dari total 5000 tweet yang diambil dan setelah melakukan penghapusan atau penghilangan terhadap tweet yang sama maka didapat hasil tweet yang sebenarnya sejumlah 1338 tweet dan melalui tahapan preprocessing, translate dan selanjutnya masuk pada tahapan metode sentiwordnet maka didapat ternyata dari total keseluruhan data tersebut setelah dianalisis ada 563 data yang positif, 751 yang negative dan ada 24 yang netral, atau dikonversikan dalam bentuk persen, yaitu 42,1 % untuk sentiment positif ,56,1% untuk sentiment negative, dan 1,8% untuk sentiment netral.

DAFTAR PUSTAKA

- [1] Media, Kemenkes, “Kementerian Kesehatan Republik Indonesia,” *kemkes.go.id*. <https://www.kemkes.go.id/article/view/20052900001/vaksin-covid-19-belum-ditemukan-pemerintah-siapkan-skenario-new-normal.html> (accessed Jul. 26, 2021).
- [2] S. Christina and D. Ronaldo, “A SURVEY OF SENTIMENT ANALYSIS USING SENTIWORDNET ON BAHASA INDONESIA,” *J. Teknol. Inf. J. Keilmuan Dan Apl. Bid. Tek. Inform.*, vol. 12, no. 2, pp. 69–73, Aug. 2018, doi: 10.47111/jti.v12i2.534.
- [3] A. Rausanfita, P. P. Adikara, and S. Adinugroho(2018), “Analisis Sentimen Twitter Menggunakan Ensemble Feature dan Metode Extreme Learning Machine (ELM) (Studi Kasus: Samsung Indonesia),” *J. Teknol. Inf.*, p. 9.
- [4] N. Yunita(2016), “ANALISIS SENTIMEN BERITA ARTIS DENGAN MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE DAN PARTICLE SWARM OPTIMIZATION,” *J. Sist. Inf.*, p. 9.
- [5] A. A. Amrullah, A. Tanton, and N. Hamdani(2016), “REVIEW ATAS ANALISIS SENTIMEN PADA TWITTER SEBAGAI REPRESENTASI OPINI PUBLIK TERHADAP BAKAL CALON PEMIMPIN,” *J. Sist. Inf.*,p. 14.
- [6] Iin Kusumawati(2017), “PROGRAM STUDI INFORMATIKA FAKULTAS KOMUNIKASI DAN INFORMATIKA UNIVERSITAS MUHAMMADIYAH SURAKARTA 2017,” *J. Sist. Inf.*, p. 16.