

Lexicon Based Sentiment Analysis pada Trending Topics di Nusa Tenggara Barat

Maria Nuzurana Asti¹, Ismarmiaty^{*2}, Ahmad Ashril Rizal³

¹Program Studi Ilmu Komputer, Fakultas Teknik, Universitas Bumigora

²Program Studi Sistem Informasi, Fakultas Teknik, Universitas Bumigora

³Program Studi Ilmu Falak, Fakultas Syariah, Universitas Islam Negeri Mataram

¹ranaedogawa@gmail.com, ^{*2}ismarmiaty@universitasbumigora.ac.id, ³ashril@uinmataram.ac.id

Abstract

This study aims to find trending topics and conduct sentiment analysis on trending topics related to the province of West Nusa Tenggara. This analysis uses tweet data from January 1, 2019 to December 31, 2020. The method used is Lexicon Based with the python programming language. The research stages consist of crawling the dataset, preprocessing, finding trending topics, lexicon based sentiment analysis & confusion matrix test and visualization. The conclusion from the analysis related to trending topics is that the top ten trending topics that emerged include: (1) COVID-19, (2) Wistan's brother, (3) MotoGP, (4) Rimpu Culture, (5) Baturotok Village fire, (6) The throwing of the factory by four housewives, (7) Rocky Gerung, (8) #onehealthKIPM, (9) after the NTB earthquake, and (10) the arrival of Sandiaga Uno to Bima. Sentiment analysis results show that several topics tend to lead to positive sentiment, including: 2nd, 3rd, 4th, 5th, 8th topics, 9th and 10th topics. While the 1st topic is related to COVID-19, the 6th and 7th tend to lead to negative sentiment. The test accuracy value is above 80% and the average score for all topics is 100% on accuracy, 100% on precision and 100% on recall.

Abstrak

Penelitian ini bertujuan untuk menemukan *trending topic* dan melakukan *sentiment analysis* terhadap *trending topic* terkait provinsi Nusa Tenggara Barat. Analisis ini menggunakan data *tweet* pada 1 Januari 2019 sampai dengan 31 Desember 2020. Metode yang digunakan adalah *Lexicon Based* dengan bahasa pemrograman *python*. Tahapan penelitian terdiri dari *crawling dataset, preprocessing, find trending topic, lexicon based sentiment analysis & confusion matrix test* dan *visualization*. Kesimpulan dari hasil analisis terkait *trending topics* adalah deretan sepuluh besar trending topik yang muncul antara lain: (1) COVID-19, (2) Adik Wistan, (3) MotoGP, (4) Budaya Rimpu, (5) kebakaran Desa Baturotok, (6) Pelemparan pabrik oleh empat ibu rumah tangga, (7) Rocky Gerung, (8) #onehealthKIPM, (9) pasca gempa bumi NTB, dan (10) Kedatangan Sandiaga Uno ke Bima. Hasil sentimen analisis didapatkan bahwa beberapa topik cenderung mengarah ke sentimen positif antara lain: topik ke-2, ke-3, ke-4, ke-5, ke-8, topik ke-9 dan ke-10. Sedangkan topik ke-1 yaitu terkait COVID-19, ke-6 dan ke-7 cenderung mengarah pada sentimen negatif. Nilai akurasi pengujian diatas 80% dan nilai rata-rata yang dihasilkan dari keseluruhan topik adalah 100% pada akurasi, 100% pada presisi dan 100% pada *recall*.

Kata kunci: *Twitter Sentiment Analysis, Lexicon Based, Trending Topik*

1. Pendahuluan

Pejabat Pengelola Informasi dan Dokumentasi (PPID) merupakan sebuah unit pengelola informasi pada Biro Humas dan Protokol Pemerintah provinsi Nusa Tenggara Barat (NTB) yang bertanggungjawab di bidang penyimpanan, pendokumentasian, penyediaan, dan/ atau pelayanan informasi badan publik sesuai amanat Undang-Undang nomor 14 tahun 2008 tentang Keterbukaan Informasi Publik [1]. Dalam rangka meningkatkan pelayanan informasi publik, pemerintah provinsi NTB melalui unit PPID berupaya menampung aspirasi masyarakat dengan dengan memanfaatkan

teknologi informasi dan sosial media dalam peningkatan pelayanan informasi publik. Penggunaan dan pemanfaatan media sosial dengan strategi yang baik dapat memberikan dampak yang maksimal terhadap pemerintah Provinsi Nusa Tenggara Barat, serta dapat mendukung pemerintah dalam menjalankan atau mencapai visi dan misinya[2].

Hasil riset yang dilakukan oleh Hootsuite (*We are Social*) yang dirilis pada bulan Januari tahun 2020 menyebutkan bahwa jumlah pengguna sosial media di Indonesia mencapai 150 juta pengguna atau sebesar 56% dari total populasi di Indonesia [3] dan meningkat pada sepanjang tahun 2020 menjadi 160 juta pengguna

atau sebesar 59% dari seluruh populasi di Indonesia. [4]. Jumlah Pengguna Indonesia yang dapat terhubung dengan iklan *Twitter* pada laporan *Twitter* adalah sebanyak 10,65 Juta Pengguna.[5] Sosial media dapat menjadi salah satu sumber data informasi yang dapat memberikan informasi dari berbagai bidang. Salah satu media social yang secara massif digunakan adalah *twitter*. *Twitter* sendiri merupakan kategori *microblogging* yang memungkinkan penggunaanya untuk mengirim dan membaca pesan yang dikenal dengan sebutan kicauan (*tweet*) [6]. Fitur utama komunikasi dalam media sosial *twitter* adalah *tweet* (kicauan) dimana pengguna dapat membuat tulisan dengan panjang karakter sebanyak 280 karakter [7]. Dari kumpulan data *tweet* terkait dengan sebuah topik dapat dilakukan analisis sentimen terhadap respon terkait topik tersebut.

Sentiment analysis merupakan salah satu *text mining* untuk mengetahui tingkat emosi dari himpunan teks yang ada. Tingkat emosi terbagi dalam dua kelas, yaitu kelas positif dan negatif. Penelitian ini bertujuan untuk mengetahui *trending topic* dan melakukan *sentiment analysis* terhadap *trending topic* terkait di provinsi Nusa Tenggara Barat. Rentang waktu data sejak 1 Januari 2019 sampai 31 Desember 2020. Metode analisis yang digunakan adalah *Lexicon Based* dengan bahasa pemrograman *python*. Dengan jumlah topik yang akan diambil sebanyak 10 topik trending. Hasil analisis penelitian diharapkan dapat dimanfaatkan sebagai rekomendasi pengambilan keputusan program, strategi pemerintah dan bahan evaluasi kinerja pemerintah.

Beberapa penelitian dengan topik yang sama pernah dilakukan adalah penelitian [8] terkait analisis sentiment tentang opini maskapai penerbangan pada dokumen *twitter* menggunakan algoritma *Support Vector Machine* (SVM), penelitian [9] dengan judul analisis sentiment topik viral desa Penari pada media sosial *twitter* dengan metode *Lexicon Based*, penelitian [10] analisis sentimen pengguna Gopay menggunakan metode *Lexicon Based* dan *Support Vector Machine* (SVM), penelitian [11] tentang analisis sentimen *twitter* pasca pengumuman hasil pilpres 2019 menggunakan metode *Lexicon analysis*, penelitian [12] terkait analisis sentimen situs pembajak artikel penelitian menggunakan metode *Lexicon-Based*, penelitian [13] terkait dengan analisis sentimen masyarakat pada media sosial *twitter* terhadap penerapan new normal di indonesia menggunakan metode *Lexicon*, penelitian [14] terkait dengan analisis sentimen pengguna *twitter* terhadap dompet elektronik dengan metode *Lexicon Based* dan *K-Nearest Neighbor*, penelitian [15] terkait dengan penerapan *Lexicon Based* untuk analisis sentimen pada *twitter* terhadap isu COVID-19, penelitian [16] terkait dengan kecenderungan tanggapan masyarakat terhadap vaksin Sinovac

berdasarkan *Lexicon Based Sentiment Analysis*, penelitian [17] terkait dengan analisis Sentimen terhadap vaksin Covid-19 di Indonesia pada *twitter* menggunakan metode *Lexicon Based* dan penelitian [18] terkait dengan analisis sentimen terhadap kebijakan pemerintah dalam menangani covid-19 dengan pendekatan *Lexicon Based*.

2. Metode Penelitian

Sentiment analysis merupakan proses untuk memahami, mengekstrak, dan mengolah data berupa dokumen teks seperti artikel atau postingan di sosial media untuk mengetahui emosi dalam sebuah kalimat apakah bersifat positif, negatif atau netral [19]. Selain untuk mengetahui emosi dalam sebuah kalimat terkait suatu hal atau peristiwa, *sentiment analysis* juga digunakan untuk mengidentifikasi kecenderungan hal yang sedang terjadi dan menjadi topik pembicaraan.[11] *Sentiment analysis* juga dapat menyatakan perasaan sedih, gembira, atau marah [6] *Lexicon based* merupakan salah satu metode yang digunakan untuk melakukan klasifikasi. Metode ini merupakan metode yang bekerja dengan mengandalkan kamus, dimana kata-kata yang terdapat pada kamus digunakan untuk melakukan identifikasi terhadap suatu kalimat [20]. Tahapan proses dari metode *lexicon based* terdiri dari *data acquisition, load dictionary, preprocessing, extract keyword,determine sentiment* [6]. *Twitter* merupakan salah satu sosial media yang sering digunakan saat ini. *Twitter* memiliki fitur-fitur menarik seperti *Direct Message (DM)* yang memungkinkan pengguna untuk dapat bertukar pesan secara pribadi, *reply tweet* yang digunakan untuk membalas atau memberi respon terhadap sebuah *tweet*, *retweet* untuk memposting kembali sebuah *tweet* dari akun lain, *mention* untuk menyebut seseorang dengan *username* yang digunakan, *hashtag* merupakan sebuah tanda untuk menandai suatu kata yang digunakan sebagai topik atau pembahasan dalam sebuah diskusi atau perbincangan agar mudah dicari dan ditemukan, trending topik memungkinkan pengguna *twitter* untuk mengetahui topik apa saja yang sedang hangat diperbincangkan oleh pengguna *twitter* [21].

Dataset yang akan digunakan pada penelitian ini adalah data *tweet* yang diperoleh dari *twitter* yang kemudian digunakan sebagai model dalam menentukan hasil sentimen. *Dataset* diambil melalui *Twitter Application Programming Interface (API)* dengan proses *crawling*. *Dataset* yang diambil berupa sekumpulan *tweet* yang dicari menggunakan *keyword* yang sudah ditentukan dengan rentang data *tweet* sejak tanggal 1 Januari 2019 sampai dengan 31 Desember 2020. Pengumpulan data *tweet* dilakukan dengan menggunakan *library python* yaitu *tweepy*. Tahapan yang dilaksanakan dalam penelitian seperti yang dapat dilihat pada bagan 1.



Bagan 1. Tahapan Analisis Sentimen Trending Topik pada Twitter

Bagan 1 yang menjelaskan tahapan dari penelitian yang dilakukan, dimana tahapan-tahapan penelitian dilakukan sebagai berikut:

1. *Crawling Dataset*: merupakan proses untuk mengumpulkan atau mengunduh data pada twitter dengan menggunakan bantuan API.
2. *Preprocessing*: dilakukan untuk mempersiapkan data untuk dapat diproses. Beberapa proses yang terdapat pada tahapan *preprocessing* antara lain: *lowercase*, *removing symbol*, *stemming*, *tokenizing* dan *bag of words*.
3. Menemukan Trending Topik: dilakukan dengan menghitung frekuensi kemunculan kata dengan menggunakan metode *n-gram*. Perhitungan frekuensi kata dilakukan dengan menggunakan *library nltk (natural language toolkit)*.
4. *Sentiment Analysis*: dilakukan menggunakan metode *text classification*. Hasil klasifikasi menjadi kelas positif dan kelas negatif. Hasil *sentiment analysis* akan disajikan dalam bentuk tabel daftar trending topik beserta jumlah frekuensi kemunculan. Selain itu penggambaran isu terkait dengan *trending topics* digambarkan dengan menggunakan *wordcloud*.
5. Pengujian: Pengukuran terhadap kinerja menggunakan *confussion matrix*. Hasil Hasil klasifikasi kelas positif dan kelas negatif dibagi menjadi 4 istilah yaitu: nilai *True Positive* (TP) merupakan jumlah data positif yang terdeteksi benar, nilai *True Negative* (TN) merupakan jumlah data negatif yang terdeteksi benar, nilai *False Positive* (FP) merupakan jumlah data negatif namun terdeteksi sebagai data positif dan nilai *False Negative* (FN) merupakan data positif yang terdeteksi sebagai data negatif. *Confussion Matrix* dapat disajikan seperti pada tabel 1.

Tabel 1. *Confussion Matrix*

| Kelas | Terklasifikasi Positif | Terklasifikasi Negatif |
|---------|------------------------------|------------------------------|
| Positif | TP (<i>True Positive</i>) | FN (<i>False Negative</i>) |
| Negatif | FP (<i>False Positive</i>) | TN (<i>True Negative</i>) |

Berdasarkan nilai 4 kelas pembagian tersebut dapat diperoleh nilai akurasi, presisi dan *recall*. Nilai akurasi merupakan nilai yang menggambarkan mengenai seberapa besar keakuratan sistem dalam mengklasifikasi keseluruhan data secara benar, nilai tersebut diperoleh menggunakan persamaan (1). Nilai presisi merupakan nilai yang menggambarkan jumlah kategori data positif yang diklasifikasi secara benar dibagi dengan total data yang terklasifikasi positif. Nilai

presisi dapat diperoleh menggunakan persamaan (2). Sedangkan *recall* menunjukkan berapa persen data kategori positif yang terklasifikasikan dengan benar. Nilai *recall* dapat diperoleh menggunakan persamaan (3).

$$\text{Akurasi} = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad (1)$$

$$\text{Presisi} = \frac{TP}{FP+TP} \times 100\% \quad (2)$$

$$\text{Recall} = \frac{TP}{FN+TP} \times 100\% \quad (3)$$

Akurasi, menggambarkan seberapa akurat model dalam mengklasifikasi dengan benar. Presisi menggambarkan akurasi anatara data yang diminta dengan hasil prediksi yang diberikan oleh model. Dan *recall* menggambarkan keberhasilan model dalam menemukan kembali sebuah informasi.

3. Hasil dan Pembahasan

Pada tahapan *crawling data* didapatkan 45.198 *tweet*. *Dataset* tersebut diolah pada tahapan *preprocessing* sehingga dapat digunakan dalam tahapan pencarian trending topik. Jumlah yang digunakan sebagai trending topik yang dibahas sebanyak 10 topik ditentukan dari jumlah maksimal kemunculan topik dengan kata kunci yang sudah ditentukan. Topik trending, sebaran dari frekuensi dan persentase kemunculan topik dapat dilihat pada diagram 1.

Frekuensi Kemunculan Topik Sebagai Trending Topik

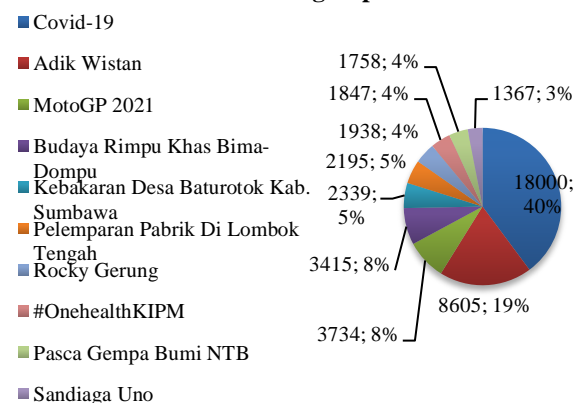
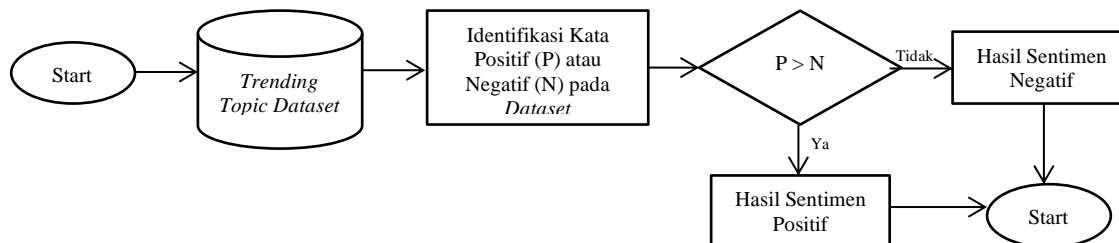


Diagram 1. Persentase Frekuensi Trending Topik NTB

Diagram 1 menunjukkan frekuensi kemunculan topik yang paling tinggi yaitu tertinggi pada topik COVID-19 dengan jumlah *tweet* sebanyak 18.000 *tweet*. Topik kedua adalah topik Adik Wistan sebanyak 8.605 *tweet* dengan disusul oleh topik ketiga yaitu MotoGP 2021 sebanyak 3.734 *tweet*. Topik terbanyak keempat adalah Budaya Rimpu Khas Bima-Dompu, lalu kelima terkait

Kebakaran Desa Baturotok Kabupaten Sumbawa dan keenam adalah topik Pelemparan Pabrik di Lombok Tengah. Sedangkan empat topik dengan jumlah kemunculan terendah yaitu terkait Rocky Gerung, gerakan #OnehealthKIPM, pasca Gempa Bumi NTB dan topik Sandiaga Uno. Kesepuluh trending topik

yang dihasilkan perlu untuk diklasifikasikan ke dalam kelas negatif dan positif dengan menggunakan *text classification* dengan metode *Lexicon Based*. Alur perhitungan metode *Lexicon Based* adalah seperti pada bagan 2.



Bagan 2. Alur Klasifikasi Data Sentimen pada *Lexicon Based Method*

Hasil yang didapatkan dari klasifikasi *dataset* ke dalam kelas positif dan negatif didapatkan hasil 29.517 data yang masuk ke dalam kelas positif dan 15.681 data ke dalam kelas negatif. Hasil klasifikasi data tersebut membagi kelas menjadi kelas *True Positif* (TF) dan *False Negatif* (FN) yang berasal dari Kelas Positif serta kelas *False Positif* (FP) dan *True Negatif* (TN) dari Kelas Negatif. Untuk mengetahui tingkat kinerja dari proses klasifikasi yang dilakukan maka pengujian dilakukan dengan menggunakan *Confussion Matrix* dimana perhitungan nilai akurasi dengan menggunakan persamaan (1), presisi dengan menggunakan persamaan (2) dan *recall* dengan menggunakan persamaan (3). Hasil klasifikasi empat kelas dan uji akurasi, presisi dan *recall* dapat dilihat di tabel 2.

Tabel 2. Hasil Pengujian *Confussion Matrix*

| No. | Trending topik | Kelas Positif | | Kelas Negatif | | Confusion Matrix | | |
|-----------------|---------------------------------------|---------------|-----|---------------|--------|----------------------|--------|--------|
| | | TP | FN | FP | TN | A | P | R |
| 1. | COVID-19 | 6.145 | 157 | 107 | 11.591 | 100,0% | 100,0% | 100,0% |
| 2. | Adik Wistan | 8.603 | 0 | 0 | 2 | 99,8% | 100,0% | 99,8% |
| 3. | MotoGP 2021 | 3.728 | 2 | 0 | 4 | 83,0% | 100,0% | 82,7% |
| 4. | Budaya Rimpu Khas Bima-Dompu | 3.409 | 6 | 0 | 0 | 100,0% | 100,0% | 100,0% |
| 5. | Kebakaran Desa Baturotok Kab. Sumbawa | 2.339 | 0 | 0 | 0 | 86,3% | 95,6% | 98,7% |
| 6. | Pelemparan Pabrik di Lombok Tengah | 133 | 21 | 6 | 2.035 | 100,0% | 100,0% | 100,0% |
| 7. | Rocky Gerung | 2 | 0 | 0 | 1.936 | 99,8% | 100,0% | 99,8% |
| 8. | #OnehealthKIPM | 1.529 | 318 | 0 | 0 | 99,9% | 100,0% | 99,9% |
| 9. | Pasca Gempa Bumi NTB | 1.755 | 3 | 0 | 0 | 100,0% | 100,0% | 100,0% |
| 10. | Sandiaga Uno | 1.367 | 0 | 0 | 0 | 97,5% | 98,2% | 98,5% |
| Jumlah = | | 29.010 | 507 | 113 | 15.568 | 100,0% | 100,0% | 100,0% |
| | | 29.517 | | 15.681 | | (%) Rata-Rata | | |



Gambar 1. Visualisasi *Wordcloud* Trending Topik 1 terkait COVID-19 dan Visualisasi *Wordcloud* Trending Topik 2 terkait Adik Wistan

Tabel 2 memberikan hasil analisis bahwa trending topik tertinggi yaitu terkait COVID-19 cenderung menuju ke arah sentimen negatif dengan jumlah sentimen pada

adalah sebanyak 11.748 data dengan nilai akurasi 100%. Visualisasi *wordcloud* untuk *trending topic* 1 dan *trending topic* 2 dapat dilihat pada gambar 1. Dari gambar 1 dapat kita lihat bahwa jumlah kata-kata yang berhubungan dengan topik 1 sangat amat banyak dan arah hubungan kata-kata mengarah ke arah negatif, sedangkan pada topik 2 sebaran kata-kata mengarah ke positif dengan kata-kata ajakan untuk membantu Adik Wistan.

Trending topik pada posisi ketiga adalah topik terkait MotoGP, topik ini cenderung menuju kearah positif dengan jumlah *tweet* sebanyak 3.728 dengan nilai akurasi 83%. Trending topik pada posisi keempat

adalah topik terkait Budaya Rimpu cenderung menuju kearah positif dengan jumlah *tweet* 3.409 dengan nilai akurasi 100%. Visualisasi *wordcloud* untuk *trending topic* 3 dan *trending topic* 4 dapat dilihat pada gambar 2.



Gambar 2. Visualisasi *Wordcloud* Trending Topik 3 terkait MotoGP2021 dan Visualisasi *Wordcloud* Trending Topik 4 terkait Budaya Rimpu

Pada gambar 2 dapat dilihat bahwa kecenderungan positif terhadap topik MotoGP2021 terlihat dari kata-kata yang berhubungan dengan antusiasme masyarakat terhadap gelaran acara *event* internasional tersebut di sirkuit Mandalika. Sedangkan pada topik Budaya Rimpu yang juga memiliki keenderungan sentimen positif dengan melihat sebaran kata-kata yang terbentuk dalam *wordcloud* berhubungan dengan kata-kata positif antara lain seperti aku cinta, cinta budaya, dan budaya nusantara.



Gambar 3. Visualisasi *Wordcloud* Trending Topik 5 terkait Kebakaran Desa Baturotok dan Visualisasi *Wordcloud* Trending Topik 6 terkait Pelemparan Pabrik Lombok Tengah

Trending topik pada posisi kelima adalah topik terkait kebakaran di Desa Baturotok Kabupaten Sumbawa, topik ini mengarah kearah positif dengan jumlah *tweet* 2.339 dengan nilai akurasi 86,3%. *Trending topic* pada posisi keenam adalah topik terkait pelemparan pabrik oleh empat ibu rumah tangga. Hasil cenderung menuju kearah negatif dengan jumlah *tweet* 2.056 dengan nilai akurasi 100%. Sedangkan, visualisasi *wordcloud* untuk *trending topic* 5 dan *trending topic* 6 dapat dilihat pada gambar 3. Pada gambar 3 dapat dilihat bahwa kecenderungan positif dari topik 5 terkait kebakaran berhubungan dengan ajakan secara positif kepada pengguna *twitter* untuk membantu korban kebakaran yang terjadi. Sedangkan kecenderungan negative yang muncul pada topic 6 terkait Pelemparan Pabrik Lombok Tengah oleh ibu-ibu rumah tangga berhubungan dengan kata-kata proses hukum dan ancaman penjara. *Trending topic* pada posisi ketujuh adalah topik terkait kedatangan Rocky Gerung dimana hasil klasifikasi cenderung menuju kearah negatif dengan jumlah *tweet*

1.936 dengan nilai akurasi 99,8%. Trending topik pada posisi kedelapan adalah topik terkait #onehealthkipm, hasil klasifikasi topik ini cenderung menuju kearah positif dengan jumlah *tweet* 1529 dengan nilai akurasi 99,9%.



Gambar 4. Visualisasi *Wordcloud* Trending Topik 7 terkait Rocky Gerung dan Visualisasi *Wordcloud* Trending Topik 8 terkait #OneHealthKIPM

Visualisasi *wordcloud* untuk *trending topic* 7 dan *trending topic* 8 dapat dilihat pada gambar 4. Pada gambar 4 tersebut dapat dilihat bahwa topik 7 terkait dengan Rocky Gerung cenderung negatif terlihat dari kata-kata yang muncul di *wordcloud* antara lain dengan kata *demo* (demonstrasi), rocky gerung gagal dan gagal *happy*. Sedangkan topik ke-8 cenderung positif berhubungan dengan dukungan masyarakat terkait pengawasan, kegiatan, perikanan, lintas perikanan dan lainnya.

Trending topic kesembilan adalah topik terkait pasca gempa bumi di Nusa Tenggara Barat, dimana hasil klasifikasi terkait topik ini cenderung menuju kearah positif dengan jumlah *tweet* 1.755 dengan nilai akurasi 100%. *Trending topic* pada urutan terakhir adalah topik terkait kedatangan Sandiaga Uno ke Bima. Hasil klasifikasi terkait dengan kedatangan Sandiaga Uno ke Bima cenderung menuju kearah positif dengan jumlah *tweet* 1.367 dengan nilai akurasi 97,5%. Visualisasi *wordcloud* untuk *trending topic* 9 dan *trending topic* 10 dapat dilihat pada gambar 5.



Gambar 5. Visualisasi *Wordcloud* Trending Topik 9 terkait Pasca Gempa Bumi NTB dan Visualisasi *Wordcloud* Trending Topik 10 Sandiaga Uno

Pada gambar 5 ditampilkan bahwa topik ke-9 terkait Pasca Gempa Bumi cenderung kearah positif dengan hubungan kata-kata di *wordcloud* berhubungan dengan kata-kata antara lain rehabilitasi, rapat terbatas, rekonstruksi dan lainnya. Topik terakhir (ke-10) yaitu terkait topik Sandiaga Uno yang mengunjungi Bima cenderung ke sentiment positif dengan melihat hubungan kata-kata yang tampil pada *wordcloud*

berhubungan dengan kata-kata antara lain antusias masyarakat, terharu, menyambut kedatangan dan lainnya.

4. Kesimpulan

Kesimpulan dari hasil analisis terkait *trending topics* adalah deretan sepuluh besar trending topik yang muncul antara lain: (1) COVID-19, (2) Adik Wistan, (3) MotoGP, (4) Budaya Rimpu, (5) kebakaran Desa Baturotok, (6) Pelemparan pabrik oleh empat ibu rumah tangga, (7) Rocky Gerung, (8) #onehealthKIPM, (9) pasca gempa bumi NTB, dan (10) Kedatangan Sandiaga Uno ke Bima. Hasil analisis sentimen didapatkan bahwa beberapa topik cenderung mengarah ke sentimen positif antara lain: topik ke-2 (Adik Wistan), topik ke-3 (MotoGP), topik ke-4 (Budaya Rimpu), topik ke-5 (Kebakaran Desa Baturotok), topik ke-8 (#onehealthKIPM), topik ke-9 (pasca gempa bumi NTB) dan topik ke-10 (kedatangan Sandiaga Uno ke Bima). Sedangkan topik ke-1 yaitu terkait COVID-19, topik ke-6 (pelemparan pabrik oleh empat ibu rumah tangga) dan topik ke-7 terkait Rocky Gerung cenderung mengarah pada sentimen negatif. Nilai akurasi pengujian diatas 80% menunjukkan nilai yang baik dan nilai rata-rata yang dihasilkan dari keseluruhan topik adalah 100% pada akurasi, 100% pada presisi dan 100% pada *recall* sehingga pengujian data dianggap dapat memberikan gambaran baik terkait dengan klasifikasi yang dilakukan.

Daftar Rujukan

- [1] Kemenkumham, *Undang Undang Republik Indonesia No 14 Tahun 2008 Tentang Keterbukaan Informasi Publik*. 2008.
- [2] M. A. Furqon, D. Hermansyah, S. Sari, A. Sukma, Y. Akbar, and N. A. Rakhmawati, "Analisis Sosial Media Pemerintah Daerah di Indonesia Berdasarkan Respons Warganet," *J. Sositologi*, vol. 17, no. 2, pp. 2–4, 2018.
- [3] We Are Social & Hootsuite, "Indonesia Digital report 2020," 2020.
- [4] We Are Social & Hootsuite, "Digital 2021," *Global Digital Insights*. p. 103, 2021.
- [5] H. Nurhayati-Wolff, "Internet usage in Indonesia - statistics & facts," *Statista*, 2021.
- [6] Stasiswaty, Rusnia, and N. Ransi, "Analisis Sentimen Wisata Bahari Di Sulawesi Tenggara Memanfaatkan Media Sosial Twitter Dengan Menggunakan Metode Lexicon," *semantik*, vol. 3, no. 2, pp. 161–168, 2017, [Online]. Available: <http://ojs.uho.ac.id/index.php/semantik/article/view/3657>.
- [7] A. Rosen, "Mencuit akan Jadi Lebih Mudah," *blog.twitter.com*, 2017.
- [8] A. M. Pravina, I. Cholissodin, and P. P. Adikara, "Analisis Sentimen Tentang Opini Maskapai Penerbangan pada Dokumen Twitter Menggunakan Algoritme Support Vector Machine (SVM)," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 3, no. 3, pp. 2789–2797, 2019, [Online]. Available: <http://j-ptiik.ub.ac.id>.
- [9] R. Arief and K. Imanuel, "Analisis Sentimen Topik Viral Desa Penari Pada Media Sosial Twitter Dengan Metode Lexicon Based," *J. Ilm. Matrik*, vol. 21, no. 3, pp. 242–250, 2019, doi: 10.33557/jurnalatrik.v21i3.727.
- [10] R. Mahendrajaya, G. A. Buntoro, and M. B. Setyawan, "Analisis Sentimen Cyberbullying pada Komentar Instagram dengan Metode SVM," *J. Tek. Univ. Muhammadiyah Ponorogo*, vol. 3, no. 2, pp. 52–63, 2019, [Online]. Available: <http://studentjournal.umpo.ac.id/index.php/komputek>.
- [11] P. A. R. Nababan, A. S. M. Lumenta, Y. D. Y. Rindengan, F. J. Pontoh, and Y. V. Akay, "Analisis Sentimen Twitter Pasca Pengumuman Hasil Pilpres 2019 Menggunakan Metode Lexicon Analysis," *J. Tek. Inform.*, vol. 15, no. 1, pp. 33–44, 2020.
- [12] M. G. Undap, V. P. Rantung, and P. T. D. Rompas, "Analisis Sentimen Situs Pembajak Artikel Penelitian Menggunakan Metode Lexicon-Based," *JOINTER – J. INFORMATICS Eng.*, vol. 2, no. 2, pp. 39–46, 2021.
- [13] A. J. Dahur and I. Albanna, "Analisis Sentimen Masyarakat Pada Media Sosial Twitter Terhadap Penerapan New Normal Di Indonesia Menggunakan Metode Lexicon Based," in *Seminar Nasional Sains dan Teknologi Terapan IX 2021*, 2020, pp. 201–206.
- [14] S. S. Salim and J. Mayary, "Analisis Sentimen Pengguna Twitter Terhadap Dompot Elektronik Dengan Metode Lexicon Based Dan K – Nearest Neighbor," *J. Ilm. Inform. Komput.*, vol. 25, no. 1, pp. 1–17, 2020, doi: 10.35760/ik.2020.v25i1.2411.
- [15] Y. N. Prasetya, D. Winarso, and Syahril, "Penerapan Lexicon Based Untuk Analisis Sentimen Pada Twitter Terhadap Isu Covid-19," *Fasilkom*, vol. 11, no. 2, pp. 97–103, 2021.
- [16] Dewi Hernikawati, "Kecenderungan Tanggapan Masyarakat Terhadap Vaksin Sinovac Berdasarkan Lexicon Based Sentiment Analysis The Trend of Public Response to Sinovac Vaccine Based on Lexicon Based Sentiment Analysis," *J. Ilmu Pengetah. dan Teknol. Komun.*, vol. 23, no. 1, pp. 21–31, 2021, [Online]. Available: <http://dx.doi.org/10.33169/jiptekkom.23.1.2021.21-31>.
- [17] P. A. Sumitro, D. I. Mulyana, and W. Saputro, "Analisis Sentimen Terhadap Vaksin Covid-19 di Indonesia pada Twitter Menggunakan Metode Lexicon Based," *J-ICOM-Jurnal Inform.*, vol. 02, no. 02, pp. 50–56, 2021, [Online]. Available: <https://www.ejurnalunsam.id/index.php/jicom/article/view/4009>.
- [18] F. S. Yerzi and Y. Sibaroni, "Analisis Sentimen Terhadap Kebijakan Pemerintah Dalam Menangani Covid-19 Dengan Pendekatan Lexicon Based," vol. 8, no. 5, pp. 11354–11366, 2021.
- [19] I. F. Rozi, S. H. Pramono, and E. A. Dahlan, "Implementasi Opinion Mining (Analisis Sentimen) untuk Ekstraksi Data Opini Publik pada Perguruan Tinggi," *J. EECIS*, vol. 6, no. 1, 2012.
- [20] Y. Azhar, "Metode Lexicon-Learning Based Untuk Identifikasi Tweet Opini Berbahasa Indonesia," *J. Nas. Pendidik. Tek. Inform.*, vol. 6, no. 3, 2017, doi: 10.23887/janapati.v6i3.11739.
- [21] A. E. P. Harrera, "Pemanfaatan Media Sosial Twitter oleh Ridwan Kamil dan Ganjar Pranowo Telah Sesuai dengan Fungsi Utama Media Massa," *J. Messenger*, vol. 8, no. 2, p. 52, 2016, doi: 10.26623/themessenger.v8i2.335.