

Sentimen Analisis Opini Masyarakat Pada Sosial Media Twitter Terhadap Organisasi Aksi Cepat Tanggap Menggunakan Naïve Bayes Classifier

¹Muhammad Fany Fahrezi, ^{2*}Angga Aditya Permana

¹Prodi Teknik Informatika, Fakultas Teknik, Universitas Muhammadiyah Tangerang, Jl. Perintis Kemerdekaan 1/33, Cikokol Kota Tangerang

²Prodi Informatika, Fakultas Teknik dan Informatika, Universitas Multimedia Nusantara, Jl. Boulevard Gading Serpong, Kelapa Dua Kab Tangerang
e-mail: angga.permana@umn.ac.id

Receive: 30 Agustus 2022

Accepted: 27 Oktober 2022

Abstract

The existence of a topic of rapid response action issues raises various kinds of responses from the community. The number of notifications that come from many sources can affect people's perceptions. Various kinds of perceptions of the quick response action experienced by the community can trigger anger and confusion. The purpose of this study is to conduct sentiment analysis using google colab with the python programming language, to find out public opinions on news about organization act with key rapid response actions in Indonesia. This research uses google colab with python programming language to find and retrieve data from twitter social media. then it will be continued with the preprocessing stage to clean the data that has been taken from twitter. then look for the percentage of the topic of the problem, namely quick response action with the naive Bayes classifier method which is classified into positive, negative, and neutral opinions. The results of this sentiment analysis will be described in the form of barplot visualization, word cloud, and in the form of UML (Unified Modeling Language) diagrams. That way this research will produce sentiment analysis information using google colab with the python programming language in retrieving and classifying data.

Keywords: Analysis Sentiment, Action Response, Naïve Bayes Classifier, Python

Abstrak

Adanya suatu topik permasalahan aksi cepat tanggap (act) menimbulkan berbagai macam tanggapan dari masyarakat. Banyaknya pemberitahuan yang berasal dari banyak sumber dapat mempengaruhi persepsi masyarakat. Berbagai macam penilaian terhadap organisasi act yang dialami masyarakat dapat menjadi pemicu terjadinya kemarahan dan kebingungan. Tujuan penelitian ini adalah untuk melakukan analisis sentimen menggunakan *google colab* dengan bahasa pemrograman *python*, untuk mengetahui opini-opini masyarakat pada media sosial *twitter* terhadap kabar mengenai organisasi act dengan kata kunci aksi cepat tanggap yang ada di Indonesia. Penelitian ini menggunakan *google colab* dengan bahasa pemrograman *python* untuk mencari dan mengambil data dari media sosial *twitter*. kemudian akan dilanjutkan dengan tahap *preprocessing* untuk membersihkan data yang telah diambil dari *twitter*. lalu mencari persentase dari topik permasalahan yaitu aksi cepat tanggap dengan metode *naïve bayes classifier* yang diklasifikasikan kedalam opini positif, negatif, dan netral. Hasil analisis sentimen ini akan dijabarkan dalam bentuk visualisasi *barplot*, *wordcloud*, dan dalam bentuk diagram UML (*Unified Modelling Language*). Dengan begitu penelitian ini akan menghasilkan sebuah informasi analisis sentimen menggunakan *google colab* dengan bahasa pemrograman *python* dalam mengambil dan mengklasifikasi data.

Kata Kunci: Sentimen Analisis, Aksi Cepat Tanggap, Naïve Bayes Classifier, Python.

PENDAHULUAN

Pada masa ini masyarakat modern sangat bergantung pada penggunaan teknologi digital, memiliki dan menggunakan media sosial. Media sosial merupakan sebuah media *online* yang penggunanya dapat dengan mudah berbagi dan menciptakan isi meliputi blog, jejaring sosial, wiki forum dan virtual. Salah satu media sosial yang banyak digunakan yaitu *Twitter* (Patil & Kolhe, 2021) (Steven & Wella, 2020). *Twitter* merupakan salah satu media sosial *microblog* yang memungkinkan penggunanya untuk menulis tentang berbagai topik dan membahas isu-isu yang sedang terjadi. *Twitter* memiliki banyak pengguna aktif sehingga akan memberikan komentar atau informasi terbaru tentang hal yang sedang diperbincangkan di dunia dan menimbulkan *trending topic twitter*.

Tingginya pengguna *twitter* menjadikan peluang untuk membagikan informasi, ide gagasan, menuangkan isi hati atau tentang hiburan. Pengguna dapat memberikan komentar atau opini tentang permasalahan atau berita yang sedang ramai dibicarakan. Informasi *tweet* tersebut akan menjadi referensi bagi pengguna *twitter* lain. Akan tetapi dengan banyaknya *tweet* yang banyak dan tersusun secara acak menyebabkan kesulitan bagi pengguna *twitter* untuk mengetahui opini positif, negatif dan netral. Banyaknya pengguna *Twitter* yang menyampaikan komentar atau opini tersebut dapat dimanfaatkan dengan cara mengklasifikasikan *tweet* sehingga diperoleh sebuah informasi. Salah satu metode yang dapat digunakan untuk menganalisis opini-opini di *tweet* adalah analisis sentiment (Destitus et al., 2020)(Maheswari et al., 2022).

Adapun sentimen yang menjadi objek penelitian ini adalah melihat opini pengguna *twitter* terhadap kata kunci "Aksi Cepat Tanggap". dalam penelitian ini, penulis melihat salah satu media sosial populer yang disebut *twitter* dan membuat model untuk mengklasifikasikan "*tweet*" menjadi sentimen positif, negatif, dan netral. Data yang didapat dari *tweet* diproses dengan *text mining* (Juanatas et al., 2019), dilanjutkan dengan mengklasifikasikan data *tweet* ke dalam tiga kelas, positif, negatif dan netral (Indira et al., 2021).

METODE PENELITIAN

Beberapa tahap untuk mendapatkan data atau informasi dalam pemecahan masalah. Tahap-tahap yang dilakukan tersebut antara lain:

1. Studi pendahuluan

Metode pengumpulan data dengan metode kepustakaan dilakukan dengan pengumpulan jurnal, literatur, paper, makalah, maupun situs internet sebagai sumber pustaka yang berkaitan dengan materi penulisan khususnya analisis sentimen menggunakan metode *naive bayes classifier* (NBC) (Steven & Wella, 2020) .

2. Pengumpulan Data

Data yang diperoleh merupakan sumber yang diambil secara langsung dari *twitter* dengan menggunakan *google colab* dengan bahasa pemrograman *python* dengan kata kunci Aksi cepat tanggap (ACT).

3. Preprocessing Data

Lalu dilakukan *preprocessing* data untuk membersihkan data tweet yang sudah dikumpulkan, ada beberapa tahap dalam *preprocessing* yaitu *case folding*, *cleansing*, *tokenizing*, *normalization*, *stopword removal*, *stemming* (Kaur & Sarabjeet, 2021).

4. Klasifikasi Data

Setelah melewati tahap *preprocessing*, data yang sudah dibersihkan akan diklasifikasi untuk mengetahui persentase dengan menggunakan metode *naïve bayes classifier*.

5. Kesimpulan dan saran

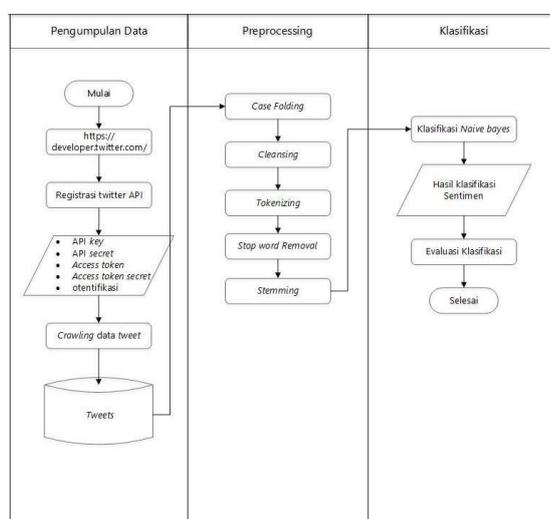
Pada tahap ini, peneliti melakukan penyimpulan dan saran terhadap hasil penelitian yang telah dilakukan. Kesimpulan ini berupa pernyataan yang diambil dari hasil dan pembahasan analisis yang telah dilakukan.

6. Penyusunan Laporan

Pada tahap ini penyusunan laporan dibuat sebagai dokumentasi yang berfungsi untuk dapat mempermudah dipelajari dan dikembangkan oleh orang lain.

HASIL DAN PEMBAHASAN

Dalam penelitian ini, berikut adalah *flowchart* analisis sentimen dimulai dari registrasi *twitter* API untuk mendapatkan data *tweet* dengan kata kunci Aksi Cepat Tanggap (ACT) hingga mendapatkan hasil klasifikasi dengan menggunakan bahasa pemrograman *python* dan metode *naïve bayes classifier* adalah sebagai berikut :



Gambar 1 Alur Analisis Sentimen

A. Pengumpulan Data

Proses pengumpulan data yang telah dilakukan berasal dari tweet yang dikirimkan oleh masyarakat di twitter. Data yang diambil berupa tweet yang menggunakan kata "Aksi Cepat Tanggap (ACT)". Data diperoleh dengan memanfaatkan *tools Python* dan *twitter* API. Sebelum melakukan *crawling* data diharuskan memiliki API *key* untuk autentifikasi dengan *Twitter*.

Proses mendapatkan API *key* :

1. Masuk ke dalam akun Twitter.
2. Masuk ke App Twitter atau buka <https://apps.twitter.com/app/new>
3. Lalu *Sign In API Development for Application on Twitter*
4. Kemudian mengisi formulir dengan informasi yang dibutuhkan.

Kemudian dapat dilakukan proses *crawling*, pertama harus dilakukan autentifikasi antara *Twitter* dengan *Python*, setelah proses autentifikasi sukses maka dapat mencari data dengan *keyword* yang diinginkan. Hasil dari *crawling* ini disimpan dalam file *csv* yang nantinya akan dilakukan tahap *preprocessing*.

B. *Pre-processing* Data

Data *Tweet* yang sudah diambil berupa data mentah yang belum siap diolah, oleh karena itu dilakukan tahap *pre-processing* untuk mendapatkan data yang siap untuk diolah pada proses selanjutnya. Tahapan *pre-processing* yang dilakukan adalah :

1. *Case Folding*

Dalam sebuah *tweet* sering kali memiliki banyak perbedaan penggunaan pada bentuk huruf, pada bagian ini dilakukan perubahan seluruh huruf kapital (*uppercase*) dikembalikan menjadi huruf kecil (*lowercase*) agar seragam.

2. *Cleansing*

Cleansing Adalah proses yang dilakukan untuk membersihkan fitur-fitur yang tidak diperlukan dalam pengambilan data yang ada pada *Twitter*, seperti *URL*, *Username*, dan lain-lain.

3. *Tokenizing*

Tokenisasi adalah proses yang dilakukan untuk memenggal kalimat menjadi beberapa bagian atau kata berdasarkan tanda bacanya seperti koma, titik, dan tanda baca lainnya

4. *Normalization*

Tahap ini dilakukan untuk mengubah kata yang tidak sesuai dengan *EYD* karena dapat mengurangi hasil sentimen, tahapan dibagi tiga langkah yaitu konversi, kata baku, dan konversi kata Bahasa Inggris.

5. *Stopword Removal*

Merupakan proses yang dilakukan untuk menghilangkan kata yang tidak diperlukan. Jika kata tersebut dibuang maka tidak akan mengubah atau menghilangkan informasi yang berada dalam kalimat tersebut. Seperti kata hubung yang, akan, di, pada dan lain-lain.

6. *Stemming*

Tahapan ini untuk membuat kata-kata yang memiliki imbuhan berubah menjadi kata dasar sesuai aturan bahasa Indonesia yang baik dan benar.

C. Hasil klasifikasi

Setelah mendapatkan hasil *translate* data *preprocessing* yang menggunakan Bahasa Indonesia ke dalam Bahasa Inggris, lalu pengklasifikasian dilakukan dengan metode *naïve bayes classification*. proses klasifikasi menentukan kata tersebut masuk kedalam klasifikasi model kata positif, negatif, dan netral, Jika dalam klasifikasi data mendapatkan score 0 itu tandanya hasil klasifikasi termasuk netral, jJika dalam klasifikasi mendapatkan score 1 itu artinya termasuk klasifikasi positif, dan jika score -1 maka artinya negatif. Klasifikasi mendapatkan data positif sebanyak 94 data, data netral sebanyak 8 data, dan data negative sebanyak 18 data.

Lalu hasil persentase klasifikasi *naïve bayes* adalah data *tweet* yang sudah melalui *preprocessing* lalu diklasifikasi dengan menggunakan *library python* yaitu *TextBlob* menghasilkan akurasi sebanyak 0.8907563025210085 atau 89%.

Selain menggunakan *naïve bayes classifier*, klasifikasi dilakukan dengan perhitungan *confusion matrix* menggunakan *library sklearn* dalam bahasa pemrograman *python*. Metode ini merupakan *tools* analisis prediktif untuk membandingkan dan menampilkan nilai sebenarnya dan nilai hasil prediksi yang menghasilkan *matriks* presisi, *recall*, *F1-Score*, dan akurasi.

Tabel dibawah ini adalah tabel *confusion matrix* otomatis dengan menggunakan *library sklearn* pada bahasa pemrograman *python* dan hasil Presisi, *Recall*, *F1-Score*, dan akurasi.

Tabel 1. Confusion Matrix

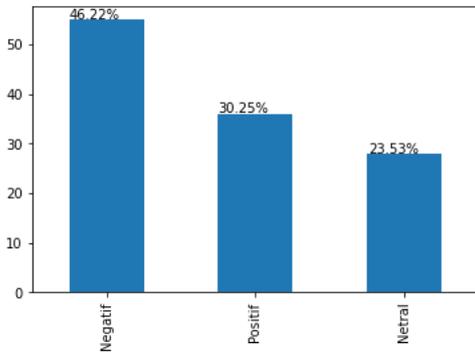
Data Prediksi				
		Positif	Netral	Negatif
Data Aktual	Positif	180	120	15
	Netral	213	148	5
	Negatif	17	115	40

Tabel 2. Hasil presisi, *recall*, *F1-Score*

Kelas	<i>Presisi</i>	<i>Recall</i>	<i>F1-Score</i>	Akurasi
Positif	0.74	0.94	0.83	0.89
Netral	1.00	0.14	0.25	
Negatif	0.93	0.94	0.93	

D. Visualisasi *Wordcloud* & *Barplot*

Hasil visualisasi *barplot* dan *wordcloud* ada pada gambar dibawah ini :



Gambar 2. Visualisasi *Barplot*



Gambar 3. Visualisasi *Wordcloud* Negatif

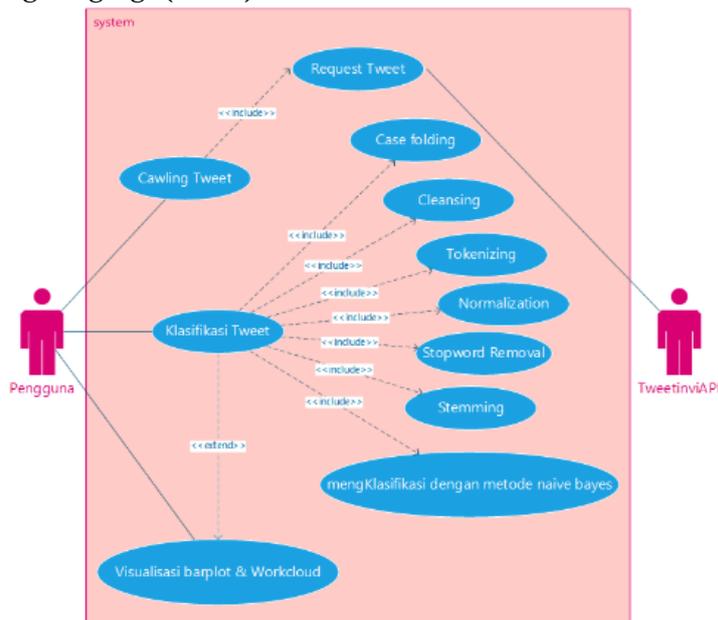


Gambar 4. Visualisasi *Wordcloud* Positif



Gambar 5. Visualisasi *Wordcloud* Netral

E. *Unified Modeling Language* (UML)



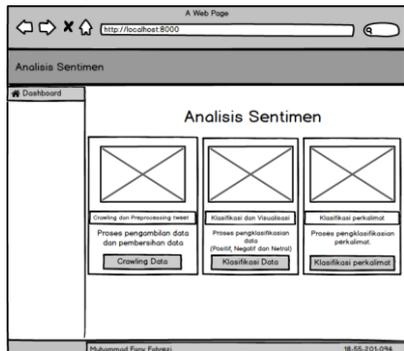
Gambar 6 Use case diagram Analisis Sentimen

Berdasarkan gambar 6 *UseCase* Diagram yang berjalan saat ini akan dijelaskan pada Tabel 3:

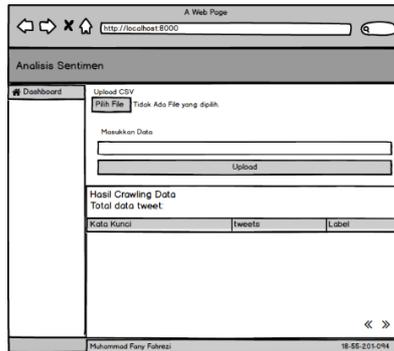
Tabel 3. Penjelasan *UseCase*

No	UseCase	Penjelasan
1	<i>Crawling Tweets</i>	Proses untuk mengambil data tweets
2	<i>Request tweets</i>	Proses untuk meminta akses data tweet pada <i>twitter API</i>
3	<i>Case folding</i>	Proses mengubah semua kata menjadi huruf kecil
4	<i>Cleansing</i>	Proses membersihkan fitur-fitur yang tidak dibutuhkan seperti, URL, username
5	<i>Tokenizing</i>	Proses memenggal kalimat menjadi beberapa bagian atau kata berdasarkan tanda baca
6	<i>Normalization</i>	Proses mengubah kata yang tidak sesuai dengan EYD
7	<i>Stopword Removal</i>	Proses menghilangkan kata yang tidak diperlukan
8	<i>Stemming</i>	Proses membuat kata-kata yang memiliki imbuhan berubah menjadi kata dasar
9	Klasifikasi tweet	Proses untuk mengklasifikasi kan tweets menjadi sentimen positif, netral, negatif
10	Mengklasifikasi dengan metode <i>naïve bayes</i>	Proses yang digunakan untuk menemukan hasil presentase sentimen dari data tweets
11	Visualisasi barplot dan <i>wordcloud</i>	Proses menampilkan hasil sentimen positif, negative, netral dalam bentuk barplot dan <i>wordcloud</i>

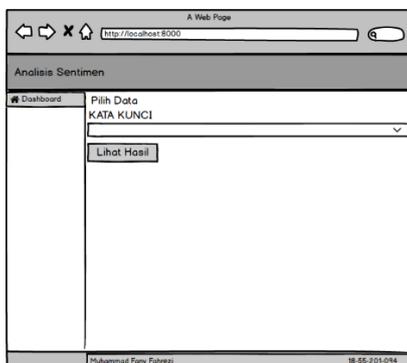
F. Perancangan Antarmuka



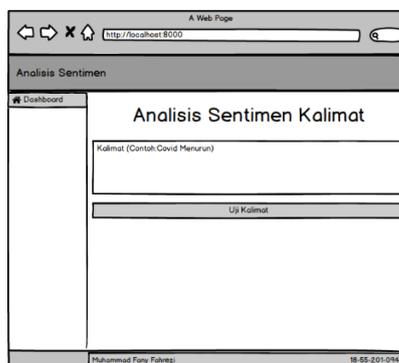
Gambar 7 Tampilan Dashboard



Gambar 8 Tampilan Menu Crawling Data



Gambar 9 Tampilan Menu Klasifikasi Data

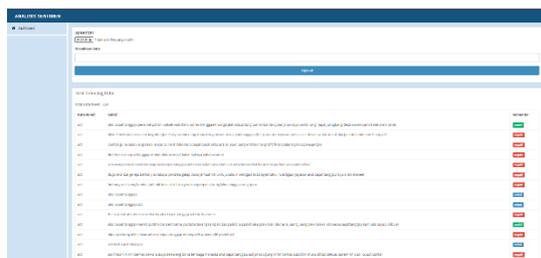


Gambar 10 Tampilan Menu Analisis Sentimen perkalimat

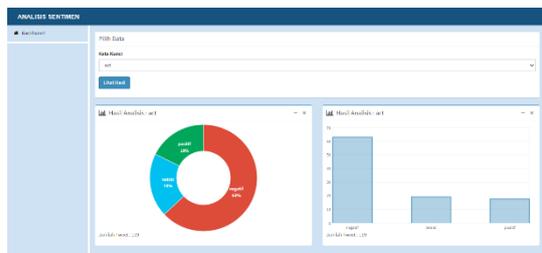
G. Implementasi Interface



Gambar 11 Halaman Dashboard



Gambar 12 Halaman Crawling Data



Gambar 13 Halaman Klasifikasi dan visualisasi



Gambar 14 Halaman Analisis Sentimen perkalimat

UCAPAN TERIMAKASIH

Penelitian ini dapat dilaksanakan dengan baik berkat bantuan dari berbagai pihak, untuk itu peneliti mengucapkan terima kasih kepada Universitas Multimedia Nusantara dan Universitas Muhammadiyah Tangerang atas dukungan dan bantuan yang telah diberikan selama proses penulisan artikel ini.

KESIMPULAN

Berdasarkan hasil penelitian, analisis sentimen opini masyarakat pada media sosial *twitter* terhadap organisasi act menggunakan metode *naïve bayes classifier (nbc)*. Dapat diperoleh kesimpulan sebagai berikut: Dengan Metode *naïve bayes classifier* dapat mengklasifikasikan sentimen pengguna *twitter* terhadap aksi cepat tanggap kedalam sentimen positif, netral, dan negatif dengan baik. Sistem Analisis menggunakan Metode *Naïve bayes classifier* menghasilkan akurasi perhitungan dengan akurasi hingga 89%.

DAFTAR PUSTAKA

- Destitus, C., Wella, W., & Suryasari, S. (2020). Support Vector Machine VS Information Gain: Analisis Sentimen Cyberbullying di Twitter Indonesia. *Ultima InfoSys : Jurnal Ilmu Sistem Informasi*, 11(2), 107–111. <https://doi.org/10.31937/si.v11i2.1740>
- Indira, D. N. V. S. L. S., Babu, C. S., Kumar, K. K., & Rao, C. V. (2021). A comprehensive review on sentiment analysis techniques and machine learning libraries in image processing. *Annals of the Romanian Society for Cell Biology*, 25(2), 4260–4267. <https://www.scopus.com/inward/record.uri?partnerID=HzOxMe3b&scp=85102993548&origin=inward>
- Juanatas, I. C., Fajardo, R. R., Manansala, E. T., Pasilan, A. A., Tabor, J. R., & Balmeo, H. D. A. (2019). Sentiment Analysis Platform of Customer Product Reviews. In *Proceedings of 2019 International Conference on Computational Intelligence and Knowledge Economy, ICCIKE 2019* (pp. 230–234). <https://doi.org/10.1109/ICCIKE47802.2019.9004415>
- Kaur, S., & Sarabjeet. (2021). Customer Segmentation Using Clustering Algorithm. In *Proceedings of International Conference on Technological Advancements and Innovations, ICTAI 2021* (pp. 224–227). <https://doi.org/10.1109/ICTAI53825.2021.9673169>
- Maheswari, T. L., Rashmi, E. I., Hasanthi, M. M., & Elakkiya, R. (2022). Customer Segmentation Based on Sentimental Analysis. In *2022 International Conference on Advanced Computing Technologies and Applications, ICACTA 2022*. <https://doi.org/10.1109/ICACTA54488.2022.9753207>
- Patil, R. S., & Kolhe, S. R. (2021). Resource Creation for Sentiment Analysis of Under-Resourced Language: Marathi. In *Communications in Computer and Information Science* (Vol. 1380, pp. 445–457). https://doi.org/10.1007/978-981-16-0507-9_37
- Steven, C., & Wella, W. (2020). The Right Sentiment Analysis Method of Indonesian Tourism in Social Media Twitter. *IJNMT (International Journal of New Media Technology)*, 7(2), 102–110. <https://doi.org/10.31937/ijnmt.v7i2.1732>