

# Analisis Text Clustering Pengguna Twitter Mengenai Statement "Indonesia Pulih Lebih Cepat Bangkit Lebih Kuat" Menggunakan Orange Data Mining

Muhamad Virji Ramadhan <sup>a,1,\*</sup>, Retno Sabrila Rahma <sup>b,2</sup>, Zalfa Putri Anzani Alimatul Ula <sup>c,3</sup>

<sup>a,b,c</sup> Universitas Nusa putra, Jln. Cibolang Cisaat – Sukabumi N0. 21, Cibolang Kaler, Kec. Cisaat, Kabupaten Sukabumi, Jawab Barat 43152

<sup>1</sup> [muhamad.virji\\_si22@nusaputra.ac.id](mailto:muhamad.virji_si22@nusaputra.ac.id); <sup>2</sup> [retno.sabrila\\_si22@nusaputra.ac.id](mailto:retno.sabrila_si22@nusaputra.ac.id); <sup>3</sup> [zalfa.putri\\_si22@nusaputra.ac.id](mailto:zalfa.putri_si22@nusaputra.ac.id)

Diterima 03 Februari 2025; Direvisi 08 Februari 2025; Diterima 10 Februari 2025

## ABSTRAK

Penelitian ini bertujuan untuk menganalisis opini pengguna Twitter terkait pernyataan "Indonesia Pulih Lebih Cepat, Bangkit Lebih Kuat" yang digaungkan pada peringatan kemerdekaan Indonesia ke-77. Analisis dilakukan menggunakan Orange Data Mining untuk mengelompokkan sentimen menjadi positif, negatif, dan netral melalui metode text clustering. Data dikumpulkan menggunakan teknik crawling melalui pustaka Python Tweepy dengan kata kunci terkait pernyataan tersebut. Setelah dilakukan preprocessing, data dianalisis menggunakan algoritma clustering seperti K-Means, serta divisualisasikan melalui scatter plot dan word cloud. Hasil analisis menunjukkan bahwa opini masyarakat mayoritas bersentimen positif terhadap tema ini, dengan dominasi emosi "surprise." Temuan ini memberikan wawasan tentang penerimaan publik terhadap kampanye pemerintah dan menunjukkan efektivitas Orange Data Mining dalam analisis teks media sosial.

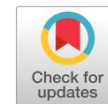


## KATA KUNCI

Analisis Sentimen;  
Text Clustering;  
Twitter;  
Orange Data Mining;  
Indonesia Pulih Lebih Cepat Bangkit;  
Lebih Kuat

## ABSTRACT

This study aims to analyze Twitter users' opinions regarding the statement "Indonesia Recovers Faster, Rises Stronger" which was echoed on the 77th anniversary of Indonesia's independence. The analysis was conducted using Orange Data Mining to categorize sentiment into positive, negative, and neutral through text clustering method. Data was collected using crawling techniques through the Python Tweepy library with keywords related to the statement. After preprocessing, the data is analyzed using clustering algorithms such as K-Means, and visualized through scatter plots and word clouds. The results of the analysis show that public opinion is predominantly positive towards this theme, with the emotion "surprise" dominating. The findings provide insights into public acceptance of the government campaign and demonstrate the effectiveness of Orange Data Mining in social media text analysis.



## KEYWORD

Sentiment Analysis;  
Text Clustering;  
Twitter;  
Orange Data Mining;  
Indonesia Recover Faster Rise;  
Stronger



This is an open-access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license

## 1. Pendahuluan

Pernyataan "Indonesia Pulih Lebih Cepat, Bangkit Lebih Kuat" yang digaungkan oleh pemerintah Indonesia berfokus pada pemulihan setelah pandemi COVID-19. Pernyataan ini mendapat respons beragam dari masyarakat, terutama di media sosial seperti Twitter, yang menjadi platform utama untuk berbagi pendapat. Oleh karena itu, penting untuk menganalisis sentimen dan pandangan masyarakat terkait pernyataan tersebut untuk memahami sejauh mana pernyataan ini diterima atau disikapi secara kritis oleh publik.

Perkembangan media sosial yang pesat telah memudahkan orang untuk melakukan berbagai tugas, termasuk bekerja, mendapatkan informasi, dan berkomunikasi dengan banyak orang di seluruh dunia. Twitter merupakan salah satu situs jejaring sosial yang memiliki keunggulan tersendiri dalam hal layanan informasi dan komunikasi[1].

Tema "Indonesia Pulih Lebih Cepat, Bangkit Lebih Kuat" diadopsi oleh Indonesia pada peringatan 77 tahun kemerdekaannya, dan dengan cepat memperoleh popularitas di Twitter. Sebagai cara untuk

meningkatkan momentum kemerdekaan, beberapa tweet muncul dengan menggunakan kata kunci “Indonesia Pulih Lebih Cepat Bangkit Lebih Kuat”.

Berdasarkan sinopsis di atas, penelitian ini mencoba menganalisis tweet dan komentar yang dibuat pengguna mengenai topik kemerdekaan, khususnya “Indonesia Pulih Lebih Cepat, Bangkit Lebih Kuat”, dengan menggunakan tiga kelas opini yang berbeda: positif, negatif, dan netral. Dengan menggunakan teks yang telah diproses sebelumnya dan algoritme pengelompokan, Orange Data Mining digunakan dalam penelitian ini.

Orange Data Mining merupakan alat yang efektif untuk analisis data teks, menggunakan berbagai metode *text clustering* untuk mengelompokkan data teks berdasarkan kemiripan konten. Melalui metode ini, kita dapat menggali topik-topik utama yang dibicarakan oleh pengguna Twitter terkait dengan pernyataan “Indonesia Pulih Lebih Cepat, Bangkit Lebih Kuat”.

## 2. Tinjauan Pustaka

### 2.1 Penelitian Terkait

Untuk memudahkan proses penelitian, penulis memanfaatkan sejumlah kajian pustaka dan tertarik dengan penelitian yang akan dilakukan. Kajian-kajian tersebut dapat digunakan sebagai sumber belajar, pembandingan, dan referensi.

Penelitian mengenai analisis sentimen dan text clustering pada media sosial telah banyak dilakukan untuk memahami persepsi publik terhadap berbagai isu. Salah satu penelitian yang relevan adalah analisis sentimen pengguna Twitter terkait kebijakan pemerintah yang dilakukan oleh Purwanto (2020). Penelitian ini menggunakan metode analisis teks untuk mengklasifikasikan tweet ke dalam sentimen positif, negatif, dan netral [4]. Studi lainnya oleh Sari et al. (2021) membahas penerapan teknik text clustering untuk mengelompokkan komentar pengguna pada platform media sosial. Tujuan utama penelitian ini adalah untuk memahami pola pikir dan opini yang dominan di kalangan pengguna media sosial [5]. Selain itu, Putra (2019) dalam penelitiannya menunjukkan efektivitas Orange Data Mining dalam mengolah data teks. Penelitian ini menyoroti fitur-fitur seperti preprocessing dan clustering yang memudahkan pengelompokan data berdasarkan kesamaan konten [6]. Penelitian lain oleh Kurniawan (2018) juga mengaplikasikan Orange Data Mining untuk analisis data teks dari media sosial. Kurniawan menekankan pada kemudahan penggunaan dan kemampuan visualisasi data yang kuat dari Orange Data Mining dalam analisis data teks[7].

### 2.2 Kerangka Berpikir

Suatu kerangka kerja ide kemudian dapat dikembangkan untuk menghasilkan solusi ideal berdasarkan landasan teori dan isu-isu yang disebutkan di atas. Untuk menawarkan solusi jangka pendek terhadap isu-isu yang diteliti, kerangka kerja konseptual pemikiran harus memiliki makna suatu konsep atau pola pikir



Gambar 1. Kerangka Berpikir

- Studi Literatur  
Untuk memahami pentingnya hubungan antara variabel, peneliti sekarang melakukan tinjauan pustaka.
- Crawling Tweet  
Memanfaatkan kumpulan data yang diperoleh dengan menggunakan Tweepy untuk mengikis data. Untuk mengumpulkan tweet dan retweet (status dan komentar), argumen pencarian memasukkan kata kunci yang terkait dengan kalimat “Indonesia Pulih Lebih Cepat, Bangkit Lebih Kuat” di kolom pencarian tweepy. Istilah pencarian #indonesiapulihlebihcepat dan #bangkitlebihkuat digunakan.

- c. Data Processing  
Transformasi, tokenisasi, normalisasi, dan penyaringan semuanya merupakan bagian dari langkah ini [2].
- d. Crowling Data Text Mining  
Data penambangan teks yang diproses sebelumnya kemudian divisualisasikan menggunakan emosi tweet dan retweet melalui prosedur visualisasi distribusi.

### 3. Metodologi Penelitian

Metodologi penelitian yang digunakan dalam pemecahan permasalahan ini mencakup metode analisis data teks melalui text clustering dengan bantuan perangkat lunak Orange Data Mining. Penelitian ini dirancang untuk mengelompokkan opini masyarakat terhadap pernyataan “Indonesia Pulih Lebih Cepat Bangkit Lebih Kuat” yang digaungkan di platform Twitter.

#### 3.1. Alat dan Bahan

1. Perangkat Lunak:
  - a. Orange Data Mining versi terbaru
  - b. Microsoft Excel atau Google Sheets untuk pengolahan data awal
2. Dataset: Data yang diambil dari Twitter menggunakan kata kunci seperti #IndonesiaPulihLebihCepat dan #BangkitLebihKuat.

#### 3.2. Pengolahan Data

Tahap pengolahan data dalam penelitian ini dilakukan melalui beberapa langkah utama. Data yang dikumpulkan dari Twitter menggunakan teknik crawling melalui pustaka Python Tweepy diproses lebih lanjut untuk memastikan kualitas dan konsistensinya.

1. Preprocessing Data: Data mentah yang diperoleh dari crawling memerlukan beberapa langkah preprocessing, seperti:
  - Tokenisasi: Membagi teks menjadi kata-kata atau simbol-simbol terpisah.
  - Normalisasi: menghapus tanda baca, simbol, dan angka dari kata-kata untuk mengembalikannya ke bentuk paling dasar.
  - Stopword Removal: Hilangkan kata-kata dari analisis yang tidak memiliki makna sebenarnya, seperti “dan,” “atau,” “yang.”
  - Filtering: Menyaring tweet yang tidak relevan dengan kata kunci penelitian seperti #IndonesiaPulihLebihCepat dan #BangkitLebihKuat.
2. Transformasi Data: Data teks yang telah diproses kemudian diubah menjadi format yang dapat digunakan untuk analisis text mining, seperti Bag of Words (BoW) atau Term Frequency-Inverse Document Frequency (TF-IDF).
3. Visualisasi Data Awal: Melakukan analisis eksplorasi data dengan memvisualisasikan distribusi sentimen awal melalui diagram batang atau word cloud untuk mendapatkan gambaran pola umum dari data yang diolah.

#### 3.3. Metode yang digunakan

Penelitian ini menggunakan pendekatan text clustering dengan memanfaatkan perangkat lunak Orange Data Mining. Tahapan metode yang digunakan yaitu:

1. Pengumpulan Data: Menggunakan teknik crawling data dengan kata kunci yang relevan di Twitter.
2. Preprocessing Data: Tahapan ini mencakup langkah-langkah yang telah dijelaskan pada bagian 3.2 untuk membersihkan dan mempersiapkan data.
3. Clustering: Data yang telah diproses dikelompokkan menggunakan algoritma clustering seperti K-Means atau Hierarchical Clustering yang tersedia di Orange Data Mining. Parameter seperti jumlah cluster ditentukan berdasarkan eksplorasi data dan evaluasi awal.
4. Evaluasi Cluster: Menggunakan metrik evaluasi seperti Silhouette Score untuk menilai kualitas pengelompokan data.

5. Visualisasi Hasil: Data cluster divisualisasikan dalam bentuk scatter plot, dendrogram, atau diagram lain yang relevan untuk memberikan wawasan intuitif terhadap pola data.

### 3.4. Jadwal Penelitian

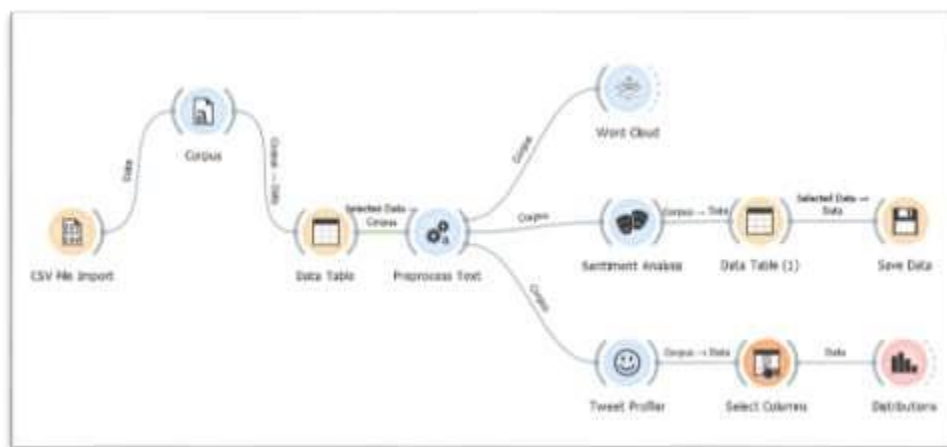
Penelitian ini direncanakan berlangsung selama empat bulan dengan tahapan yang terstruktur. Pada bulan pertama, kegiatan diawali dengan studi literatur untuk memahami variabel yang diteliti dan mengidentifikasi metode yang sesuai. Pada minggu ketiga hingga keenam, dilakukan proses crawling data dari Twitter menggunakan pustaka Python Tweepy, diikuti dengan preprocessing data seperti tokenisasi, normalisasi, dan filtering untuk memastikan data siap dianalisis. Bulan kedua dan ketiga difokuskan pada analisis data menggunakan teknik clustering dengan bantuan perangkat lunak Orange Data Mining, yang mencakup eksplorasi algoritma dan evaluasi hasil pengelompokan. Pada bulan terakhir, hasil analisis divisualisasikan dan dievaluasi lebih lanjut, diikuti dengan penyusunan laporan akhir untuk mendokumentasikan temuan penelitian secara menyeluruh.

## 4. Hasil dan Pembahasan

### 4.1 Hasil

#### a. Proses Data Mining

Pada tahap ini, desain widget text clustering ditampilkan menggunakan Orange Data Mining.



Gambar 2. Widget Text Clustering

Berdasarkan objek kajian, kumpulan data Twitter dimasukkan dan diperiksa secara menyeluruh. Untuk menciptakan visualisasi yang tepat, widget yang relevan dihubungkan berdasarkan kebutuhan analisis.

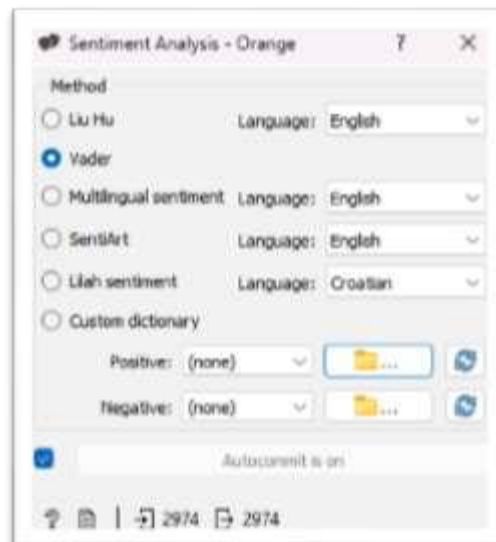
- a. **Penarikan Data (Twitter).** Kumpulan data penelitian ini terdiri dari tweet dan retweet yang berisi pembaruan status pengguna dan komentar dengan tagar #IndonesiaPulihMoreCepat dan #BangkitMore. Informasi dikumpulkan dari unggahan dalam 12 bulan terakhir. Istilah dominan yang muncul ditentukan dengan menganalisis dan meringkas temuan data. Proses memasukkan data ke dalam Orange Data Mining disebut sebagai korpus. Korpus ini memungkinkan analisis lebih lanjut, termasuk pemilihan karakteristik yang akan digunakan dalam proses analisis dan penyajian dokumen dengan jumlah baris kalimat.
- b. **Preprocess text.** Sebelum melakukan analisis teks, teks tersebut diproses terlebih dahulu, di mana teks akan ditransformasi, ditokenisasi, dinormalisasi, dan difilter setelah dibagi menjadi bagian-bagian yang lebih kecil (token). Saat ini, widget persiapan teks Orange Data Mining memungkinkan pengguna untuk mengaktifkan dan menonaktifkan fase-fase analisis yang dijalankan secara berurutan. Persiapan teks melibatkan beberapa langkah, seperti:
- c. **Transformation.** Langkah awal dalam persiapan teks adalah transformasi, yang melibatkan konversi data yang masuk menjadi huruf kecil secara default.
- d. **Tokenization.** Pada proses *tokenization*, teks akan dipecah menjadi komponen yang lebih kecil dan membagi teks berdasarkan kata, namun tidak menghilangkan simbol tanda baca.

- e. **Normalization.** *Normalization* untuk mengevaluasi dan memilih kata-kata yang akan membentuk sebuah frasa. Selain itu, metode ini dapat mengidentifikasi kesalahan ketik atau kesalahan kata dalam teks untuk menentukan maknanya.
- f. **Filtering.** *Filtering* dimaksudkan untuk menyimpan pilihan kata tertentu yang tidak diperlukan untuk langkah berikutnya atau penghapusan filter. Widget awan kata Orange Data Mining menampilkan semua data yang dibuat sebagai konsekuensi dari praproses teks. Gambar berikut menampilkan widget awan kata:



### Gambar 3. Widget Word Cloud

- g. Sentiment Analysis. Algoritma Vader, yang membagi sentimen menjadi tiga kategori positif, negatif, dan netral digunakan untuk menganalisis sentimen.



#### Gambar 4. Widget Sentiment Analysis

Vader akan mengklasifikasikan dan menetapkan skor teks dalam proses analisis sentimen berdasarkan nilai setiap kata dalam leksikonnya. Hasil akhirnya akan berupa skor gabungan total, yang akan dikompilasi dan hasilnya akan dibandingkan.

- ### 1) Data Table

Semua data yang tersedia ditampilkan menggunakan widget tabel data berdasarkan kumpulan data yang dimasukkan. Bersama dengan metode perhitungan untuk format data numerik, skor total yang juga dikenal

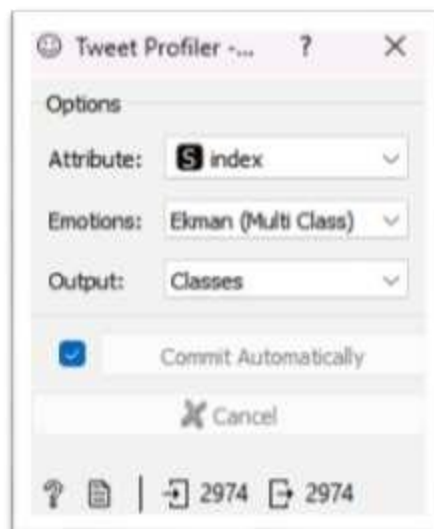
sebagai kombinasi akan digunakan untuk menghasilkan temuan analisis sentimen dan analisis teks dalam tabel data, yang akan dikategorikan ke dalam kelompok positif, negatif, dan netral.

|    | Text                | Location            | ID               | Datetime           | pos   | neg   | neu   | compound |
|----|---------------------|---------------------|------------------|--------------------|-------|-------|-------|----------|
| 1  | Aku tau gaya...     | ?                   | 1562031387527... | 2022-10-17 15:1... | 0     | 0     | 1     | 0        |
| 2  | Aku tau insang...   | ?                   | 1562031337421... | 2022-10-17 15:1... | 0     | 0     | 1     | 0        |
| 3  | untuk tahu info...  | Casuar, Jember...   | 1562031070919... | 2022-10-17 15:2... | 0     | 0     | 1     | 0        |
| 4  | RT @AntiCekro...    | Maluku, Indones...  | 1562031036497... | 2022-10-17 15:2... | 0     | 0     | 1     | 0        |
| 5  | Mau di Pasa Ma...   | ?                   | 1562030762961... | 2022-10-17 15:2... | 0     | 0     | 1     | 0        |
| 6  | RT @AntiCekro...    | ?                   | 1562029928942... | 2022-10-17 15:2... | 0     | 0     | 1     | 0        |
| 7  | Udah... Ga usah...  | Surabaya            | 1562028878912... | 2022-10-17 15:2... | 0     | 0     | 1     | 0        |
| 8  | Udah... Ga usah...  | Gorontalo           | 1562028629491... | 2022-10-17 15:1... | 0     | 0     | 1     | 0        |
| 9  | Udah... Ga usah...  | ?                   | 1562028771898... | 2022-10-17 15:1... | 0     | 0     | 1     | 0        |
| 10 | RT @Gendras S...    | ?                   | 1562027805286... | 2022-10-17 15:1... | 0     | 0     | 1     | 0        |
| 11 | RT @TVBisnis...     | Pekalongan, Ka...   | 1562027643775... | 2022-10-17 15:1... | 0     | 0     | 1     | 0        |
| 12 | Presiden Joko ...   | DKI Jakarta, Ind... | 1562027254808... | 2022-10-17 15:1... | 0     | 0     | 1     | 0        |
| 13 | 2023-Prasasti7 m... | Jakarta Pusat, D... | 1562027178915... | 2022-10-17 15:1... | 0     | 0     | 1     | 0        |
| 14 | Dipukul di km...    | Jakarta Utara, D... | 1562027158658... | 2022-10-17 15:1... | 0     | 0.104 | 0.896 | -0.1027  |
| 15 | Dipukul di km...    | Jakarta Utara, D... | 1562027152391... | 2022-10-17 15:1... | 0     | 0.104 | 0.896 | -0.1027  |
| 16 | Terlihat Terse...   | Jakarta - Indon...  | 1562027087751... | 2022-10-17 15:1... | 0.176 | 0     | 0.824 | 0.4188   |
| 17 | Aku tidak tahan...  | ?                   | 1562026801548... | 2022-10-17 15:1... | 0     | 0     | 1     | 0        |
| 18 | Benar saja kaku...  | ?                   | 1562026671815... | 2022-10-17 15:1... | 0     | 0     | 1     | 0        |
| 19 | RT @AntiCekro...    | Jakarta Capital...  | 1562025983132... | 2022-10-17 15:0... | 0     | 0     | 1     | 0        |
| 20 | Produk Baru ...     | ?                   | 1562024412282... | 2022-10-17 15:0... | 0     | 0     | 1     | 0        |
| 21 | Aku pikir saat s... | Indonesia           | 1562024058174... | 2022-10-17 15:0... | 0     | 0     | 1     | 0        |
| 22 | Buat Pemas yan...   | ?                   | 1562023890388... | 2022-10-17 15:0... | 0     | 0     | 1     | 0        |
| 23 | Buat Pemas/Pt...    | ?                   | 1562023585224... | 2022-10-17 14:5... | 0     | 0     | 1     | 0        |
| 24 | Kamik yang su...    | ?                   | 1562023152242... | 2022-10-17 14:5... | 0     | 0     | 1     | 0        |
| 25 | Kapal KRI (Ha...    | Pulau Bawean        | 1562023086967... | 2022-10-17 14:5... | 0     | 0     | 1     | 0        |
| 26 | RT @Gendras S...    | Nar peccat          | 1562022994338... | 2022-10-17 14:5... | 0     | 0     | 1     | 0        |
| 27 | Udah... Ga usah...  | ?                   | 1562022852515... | 2022-10-17 14:5... | 0     | 0     | 1     | 0        |
| 28 | Yuk join beres...   | BBM 256a722...      | 1562022989579... | 2022-10-17 14:5... | 0.128 | 0     | 0.872 | 0.296    |
| 29 | RT @AntiCekro...    | urang berant...     | 1562021647967... | 2022-10-17 14:5... | 0     | 0     | 1     | 0        |
| 30 | Temun MakiMak...    | Kiau, Indonesia     | 1562021568291... | 2022-10-17 14:5... | 0     | 0     | 1     | 0        |

Gambar 5. Widget Data Table

## 2) Tweet Profiler

Dengan menggunakan metode tweet profiler, data sentimen server untuk setiap tweet dan retweet akan dikumpulkan. Tiga klasifikasi emosi ekman, plutchik, dan profil keadaan suasana hati dikirim oleh widget ke server, yang menghitung kemungkinan dan skor emosi. Klasifikasi emosi Ekman (multikelas), analisis konten atribut, dan pengamatan variabel emosional yang direpresentasikan melalui Orange Data Mining semuanya digunakan dalam pekerjaan ini. Widget tweet profiler kemudian ditautkan ke kumpulan data yang dijelajahi menggunakan widget korpus, seperti yang terlihat pada ilustrasi di bawah ini.

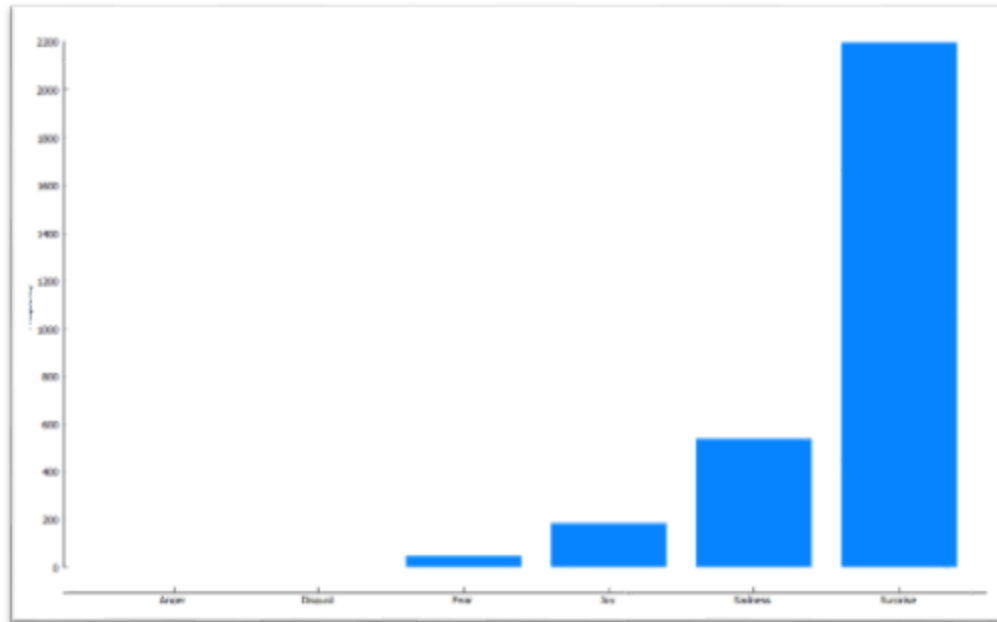


Gambar 6. Widget Tweet Profiler



### 3) Distribution

Dengan menggunakan sejumlah prosedur, distribusi nilai atribut diskrit atau kontinu ditampilkan dalam penelitian ini menggunakan widget distribusi. Enam jenis emosi dari kumpulan data tweet dan retweet yang diberikan ditampilkan oleh distribusi tersebut. Temuan distribusi tersebut menunjukkan bahwa reaksi pengguna Twitter terhadap tagar #indonesiapulihcepat dan #bangkitcepatkuat selama 12 bulan sebelumnya tidak terduga.



Gambar 7. Widget Distribution

## 4.2 Pembahasan

Temuan penelitian dan pembahasan ilmiah terkait disertakan dalam bagian hasil dan pembahasan. Tuliskan simpulan ilmiah yang ditarik dari penelitian yang telah selesai yang harus didukung oleh fakta yang memadai. Data penelitian yang diperoleh tidak sama dengan simpulan ilmiah yang dibahas di sini. Perlu untuk memberikan penjelasan ilmiah tentang temuan ini, termasuk: Apa saja temuan ilmiah yang diperoleh? Apa yang menyebabkannya terjadi? Untuk alasan apa variabel tren seperti itu? Jika diperlukan, semua pertanyaan ini harus didukung oleh fenomena dasar ilmiah yang memadai dan dijelaskan secara ilmiah daripada sekadar deskriptif. Lebih lanjut, perbandingan dengan temuan peneliti lain pada subjek yang hampir sama harus dijelaskan. Hipotesis penelitian dalam pendahuluan harus ditangani oleh temuan dan hasil penelitian.

## 5. Penutup

### 5.1. Kesimpulan

Penelitian ini menganalisis opini pengguna Twitter terkait pernyataan "Indonesia Pulih Lebih Cepat, Bangkit Lebih Kuat" menggunakan Orange Data Mining. Pernyataan tersebut, yang digaungkan pada peringatan kemerdekaan Indonesia ke-77, berhasil menjadi trending di Twitter, mencerminkan tingginya interaksi publik terhadap tema tersebut. Data yang dikumpulkan menggunakan teknik crawling melalui pustaka Python Tweepy dianalisis secara mendalam melalui preprocessing data, clustering, dan visualisasi. Hasil penelitian menunjukkan bahwa mayoritas opini masyarakat terhadap pernyataan ini memiliki sentimen positif, mengindikasikan penerimaan yang baik terhadap pesan yang disampaikan. Sentimen netral dan negatif juga ditemukan, meskipun dalam jumlah yang lebih kecil. Sentimen positif yang dominan didukung oleh emosi "surprise," yang menggambarkan antusiasme dan rasa bangga masyarakat terhadap tema kemerdekaan ini. Metode text clustering yang diterapkan menggunakan algoritma seperti K-Means membuktikan efektivitasnya dalam mengelompokkan data berdasarkan kemiripan isi. Penggunaan Orange Data Mining tidak hanya mempermudah analisis tetapi juga menyediakan visualisasi data yang intuitif, seperti scatter plot, word cloud, dan dendrogram. Dengan pendekatan ini, penelitian mampu menggambarkan pola sentimen publik secara jelas dan memberikan wawasan yang relevan tentang persepsi masyarakat terhadap kampanye pemerintah. Jika semua hal

dipertimbangkan, studi ini secara efektif menggambarkan bagaimana analisis media sosial dapat menjadi instrumen yang berguna untuk memahami bagaimana masyarakat umum bereaksi terhadap topik tertentu. Hasil ini dapat menjadi dasar bagi pemerintah dalam merancang kampanye publik yang lebih efektif di masa mendatang.

## 5.2. Saran

Penelitian selanjutnya disarankan untuk memperluas ruang lingkup analisis ke berbagai platform media sosial lainnya, seperti Instagram, Facebook, atau TikTok. Setiap platform memiliki karakteristik pengguna yang berbeda, yang dapat memberikan wawasan yang lebih kaya tentang persepsi masyarakat secara keseluruhan. Penggunaan Algoritma yang Lebih Kompleks Untuk meningkatkan akurasi analisis sentimen, disarankan untuk menggunakan algoritma yang lebih kompleks, seperti DBSCAN, LDA (Latent Dirichlet Allocation), atau metode berbasis deep learning seperti BERT (Bidirectional Encoder Representations from Transformers). Algoritma ini dapat memberikan pengelompokan data yang lebih rinci dan relevan Analisis Temporal Dianjurkan untuk melakukan penelitian dengan cakupan waktu yang lebih panjang, misalnya selama beberapa tahun, untuk mempelajari bagaimana sentimen publik berubah seiring waktu. Analisis temporal ini dapat memberikan gambaran yang lebih dinamis tentang pola opini masyarakat. Melibatkan variabel tambahan, seperti usia, lokasi, pekerjaan, atau tingkat pendidikan pengguna, akan memberikan konteks yang lebih mendalam. Variabel-variabel ini dapat membantu memahami perbedaan opini berdasarkan latar belakang demografi dan budaya masyarakat. Pemerintah, lembaga penelitian, atau praktisi media dapat memanfaatkan hasil penelitian ini untuk meningkatkan kualitas kampanye publik. Misalnya, dengan memahami respon publik secara lebih mendalam, strategi komunikasi dapat disesuaikan untuk menjangkau berbagai lapisan masyarakat secara efektif. Disarankan untuk memanfaatkan metode visualisasi data yang lebih interaktif, seperti dashboard berbasis web, untuk mempermudah pemahaman hasil analisis oleh pemangku kepentingan. Hal ini akan memberikan dampak yang lebih besar dalam pengambilan keputusan berbasis data. Penelitian selanjutnya perlu memperhatikan validasi data dengan mengintegrasikan hasil analisis dengan survei langsung kepada masyarakat. Pendekatan ini dapat meningkatkan akurasi dan memastikan bahwa hasil yang diperoleh mencerminkan kondisi sebenarnya di lapangan. Dengan berbagai saran ini, diharapkan penelitian ke depan dapat memberikan wawasan yang lebih mendalam dan relevan, tidak hanya dalam konteks analisis media sosial tetapi juga dalam mendukung strategi komunikasi yang lebih efektif di berbagai sektor.

## Daftar Pustaka

- [1] Sentiya, A., & Suroyo, H. (2019). Analisis Text Clustering Akun Fanpage Shopee Indonesia Dengan Komentar Followers Menggunakan Tools Orange Data Mining. In *Bina Darma Conference on Computer Science (BDCCS)* (Vol. 1, No. 4, pp. 1055-1067).
- [2] Nurhafida, S. I., & Sembiring, F. (2021, September). Analisis Text Clustering Masyarakat
- [3] Di Twitter Mengenai Mcdonald'Sxbts Menggunakan Orange Data Mining. In *Seminar Nasional Sistem Informasi dan Manajemen Informatika Universitas Nusa Putra* (Vol. 1, No. 01, pp. 28-35).
- [4] Negara, E. S., Andryani, R., & Saksono, P. H. (2016). Analisis data twitter: Ekstraksi dan analisis data geospasial. *INKOM Journal*, 10(1), 27-36.
- [5] Purwanto, A. (2020). Analisis Sentimen Pengguna Twitter Terkait Kebijakan Pemerintah. *Jurnal Analisis Sosial*, 15(2), 123-135.
- [6] Sari, M., & Budiman, A. (2021). Penerapan Teknik Text Clustering untuk Analisis Media Sosial. *Jurnal Teknologi Informasi dan Komunikasi*, 10(1), 45-60.
- [7] Putra, B. (2019). Efektivitas Orange Data Mining dalam Pengolahan Data Teks. *Jurnal Data Mining Indonesia*, 8(3), 76-89.
- [8] Kurniawan, D. (2018). Aplikasi Orange Data Mining untuk Analisis Data Teks Media Sosial. *Jurnal Informatika*, 7(4), 112-125.