

## ANALISIS SENTIMEN DAN AKTOR BERPENGARUH DI KOMENTAR VIDEO YOUTUBE RAYMOND CHIN MENGGUNAKAN METODE LEXICON *SENTIMENT ANALYSIS AND INFLUENTIAL ACTOR IDENTIFICATION IN RAYMOND CHIN'S YOUTUBE COMMENTS USING A LEXICON-BASED APPROACH*

Muhammad Fathan Annaba<sup>1)</sup>, Muhammad Arif Rizqi<sup>2)</sup>, Zara Zyasky Zalsabilla<sup>3)</sup>,  
Adinda Nazwa Nabila<sup>4)</sup>, Zhafirah Aqilah Izzaturahman<sup>5)</sup>

<sup>(1,2)</sup>Program Studi Ilmu Komunikasi, Fakultas Ilmu Sosial & Politik, Universitas

Pembangunan Nasional “Veteran” Jakarta, Jl. R.S Fatmawati No. 1, Cilandak, Jakarta Selatan, 12450

<sup>(3,4,5)</sup>Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional “Veteran” Jakarta, Jl. R.S Fatmawati No. 1, Cilandak, Jakarta Selatan, 12450

### Abstract

Social media has become a technological innovation today and has become part of most Indonesian people's lives. The presence of social media changes the perspective of society and the ease of access to information. One of the popular social media platforms used in Indonesia and even in the world is YouTube. In early 2024, there were 139 million YouTube users in Indonesia. Not long ago there was an issue about Tapera (Public Housing Savings) proposed by the government with the aim of helping people have permanent housing. Tapera carries the concept of mutual cooperation between communities. Of course, many reactions were generated from this issue, one of the content creators, Raymond Chin, discussed Tapera through his YouTube channel and produced various kinds of public comments. Based on these comments, sentiments can be taken to see how people respond to the Tapera policy, whether they tend to be positive, neutral, or negative. The method used is the Lexicon method, an approach in text analysis that focuses on the use of vocabulary or word dictionaries (lexicon) to understand or analyze text based on meaning or sentiment. The results were 497 positive comments, 607 neutral comments, and 699 negative comments. This study also looks at the influential actors displayed by @perjalananwaktu7181 who got 1566 likes with negative comments indicating the dominance of negative sentiment on the Tapera program.

### Article History

*Submitted: 4 Desember 2024*

*Accepted: 10 Desember 2024*

*Published: 11 Desember 2024*

### Key Words

Social media, YouTube, Sentiment analysis, Lexicon, Tapera.

### Abstrak

Media sosial telah menjadi inovasi teknologi pada saat ini dan telah menjadi bagian dari sebagian besar kehidupan masyarakat Indonesia. Hadirnya media sosial mengubah cara pandang masyarakat dan kemudahan akses dalam mendapatkan informasi. Salah satu platform media sosial yang populer digunakan di Indonesia bahkan di dunia yakni YouTube. Pada awal tahun 2024 tercatat 139 juta pengguna YouTube di Indonesia. Belum lama ini terdapat isu tentang Tapera (Tabungan Perumahan Rakyat) yang diusulkan pemerintah dengan tujuan untuk membantu masyarakat memiliki hunian tetap. Tapera mengusung konsep gotong royong antar masyarakat. Tentu banyak reaksi yang ditimbulkan dari isu ini, salah satu konten kreator yakni Raymond Chin membahas Tapera melalui channel YouTube miliknya dan menghasilkan berbagai macam komentar masyarakat. Berdasarkan komentar tersebut dapat diambil sentimennya untuk melihat bagaimana masyarakat menanggapi kebijakan Tapera, apakah cenderung positif, netral, atau negatif. Metode yang digunakan yakni metode Lexicon, salah satu pendekatan dalam analisis teks yang berfokus pada penggunaan kosakata atau kamus kata (lexicon) untuk memahami atau menganalisis teks berdasarkan arti atau sentimen. Hasilnya sebanyak 497 komentar positif, 607 komentar netral, dan 699 komentar negatif. Pada penelitian ini juga melihat aktor berpengaruh yang ditampilkan oleh @perjalananwaktu7181 yang mendapatkan 1566 likes dengan komentar negatifnya yang menandakan dominasi sentimen negatif pada program Tapera.

### Sejarah Artikel

*Submitted: 4 Desember 2024*

*Accepted: 10 Desember 2024*

*Published: 11 Desember 2024*

### Kata Kunci

Media sosial, YouTube, Analisis sentimen, Lexicon, Tapera.

## I. PENDAHULUAN

Di era digital *Society 5.0*, inovasi teknologi telah menghadirkan berbagai solusi holistik untuk mengatasi tantangan sosial dan lingkungan, khususnya melalui platform media sosial. Perkembangan teknologi yang pesat telah mengubah media sosial menjadi ruang utama untuk interaksi sosial digital. Media sosial memungkinkan akses mudah melalui smartphone atau perangkat digital lain yang terhubung dengan internet. Secara umum, media sosial adalah ruang untuk bersosialisasi, mengekspresikan diri, berkolaborasi, dan berkomunikasi dengan pengguna lain, menciptakan jejaring sosial dalam ranah digital (Nasrullah, 2015). Kehadiran media sosial ini telah mengubah cara masyarakat berkomunikasi sekaligus memberikan dampak signifikan terhadap berbagai aspek kehidupan masyarakat, terutama dalam mendukung arus komunikasi yang cepat dan dinamis.

Salah satu platform media sosial yang saat ini populer adalah *YouTube*. Sebagai platform berbasis video, *YouTube* memungkinkan penggunanya untuk berbagi konten, memberikan komentar, dan berinteraksi secara digital. Menurut laporan *We Are Social* dan *Meltwater* pada Januari 2024, *YouTube* menempati peringkat kedua sebagai media sosial yang paling banyak digunakan di Indonesia diukur melalui banyaknya waktu yang digunakan, dengan lebih dari 139 juta pengguna aktif. Popularitas *YouTube* menjadikannya ruang diskusi yang inklusif dan bebas, di mana berbagai isu sosial, ekonomi, hingga kebijakan publik dapat dibahas secara terbuka melalui kolom komentar pada video. Salah satu isu yang belakangan ramai diperbincangkan pada platform ini adalah *Tabungan Perumahan Rakyat* (Tapera), yang menjadi sorotan dalam salah satu video *YouTube* Raymond Chin.

Tapera adalah program pemerintah yang bertujuan membantu masyarakat dalam pembiayaan rumah, yang kian mahal setiap tahunnya. Program ini memiliki landasan hukum melalui *UU No. 24 Tahun 2016* dan diperbarui pada *PP No. 21 Tahun 2024*, yang menjelaskan potongan sebesar 3% dari gaji pekerja, dengan rincian 2,5% menjadi tanggungan pekerja dan 0,5% ditanggung perusahaan. Program ini wajib diikuti oleh ASN, anggota TNI/Polri, pekerja BUMN, karyawan swasta, hingga pekerja lepas, dengan syarat tertentu, seperti berusia minimal 20 tahun, memiliki gaji setara UMR, dan belum memiliki rumah pertama. Manfaat utama Tapera meliputi kemudahan pengajuan kredit, bunga rendah, dan subsidi silang, dengan sistem gotong royong sebagai fondasi utama.

Namun, isu ini menuai beragam reaksi dari masyarakat, mulai dari dukungan hingga kritik tajam, yang terpantau dalam kolom komentar video Raymond Chin. Melalui videonya, Raymond Chin membahas berbagai aspek Tapera, yang kemudian memicu diskusi aktif di kolom komentar. Reaksi publik yang terekam di kolom komentar ini mencerminkan opini masyarakat terhadap kebijakan tersebut, baik yang bersifat positif maupun negatif. Kolom komentar *Youtube* menjadi ruang terbuka, di mana pengguna menyuarakan pendapat mereka, memberikan apresiasi, atau bahkan menyampaikan kritik terkait kebijakan pemerintah.

Di tengah derasnya arus informasi ini, analisis sentimen publik terhadap komentar video Raymond Chin mengenai Tapera menjadi penting untuk memahami persepsi masyarakat. Analisis sentimen memberikan gambaran mengenai tingkat kepuasan publik, yang dapat mempengaruhi pengambilan keputusan strategis oleh pemangku kepentingan. Selain itu, keberadaan aktor berpengaruh dalam diskusi ini, yang dapat memengaruhi arah opini publik melalui tingkat *engagement* seperti jumlah *like*, *reply*, atau *mention*, juga perlu diidentifikasi. Aktor-aktor ini memiliki peran signifikan dalam membentuk persepsi masyarakat, baik mendukung maupun menentang isu tapera.

Tujuan penelitian ini adalah menganalisis sentimen publik terhadap kebijakan tapera melalui komentar video Raymond Chin di *Youtube* dan mengevaluasi aktor-aktor berpengaruh berdasarkan tingkat *engagement* mereka. Dengan memahami pola sentimen dan identifikasi aktor utama, penelitian ini diharapkan dapat memberikan *insight* bagi pemerintah dan *stakeholder* terkait dalam mengambil keputusan yang lebih informatif dan responsif terhadap

opini publik.

## II. METODE PENELITIAN

### A. Studi Literatur

Tahapan ini studi literatur ialah langkah mendasar yang menjadi fondasi serta memiliki tujuan untuk mengumpulkan, menganalisis, dan merangkum berbagai literatur relevan yang menjadi landasan teoritis dalam penelitian ini. Kajian literatur dilakukan dengan menyisir artikel jurnal, serta sumber daring kredibel yang berkaitan dengan analisis sentimen, *Lexicon-Based algorithm*, *Gap Analysis*, dan dinamika interaksi digital di media sosial, khususnya pada aplikasi *Youtube*. Studi literatur pertama pada jurnal berjudul “Analisis Sentimen Komentar YouTube tentang Terpilihnya Menteri Kabinet Indonesia Maju Menggunakan Lexicon Based” yang ditulis oleh Gupta Sanjaya dan Kemas Muslim Lhaksmana. Pada jurnal tersebut menjelaskan proses analisis sentimen pada komentar YouTube tiga channel berita besar yang menayangkan pelantikan Menteri Kabinet Indonesia Maju. Hasilnya sebanyak 2377 data komentar positif, 2261 data komentar netral, dan 725 merupakan komentar negatif. Pada jurnal tersebut juga membahas ekstraksi fitur yang digunakan lexicon Indonesia lebih baik daripada SentiWordNet.

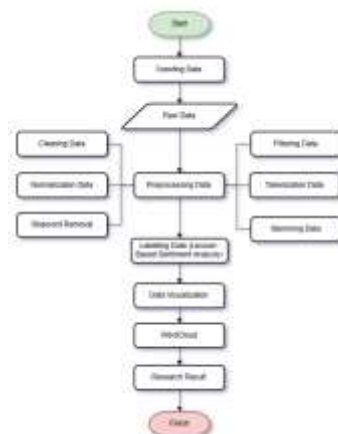
Jurnal kedua berjudul “Analisis Sentimen Pada Channel Autonetmagz Terhadap Review Mobil Almaz 2019 Dengan Metode Naive Bayes Classifier dan Lexicon Based” yang ditulis oleh Heru Prasetyo, Ghulam Asrofi Buntoro, dan Dyah Mustikasari. Pembahasan pada jurnal ini menganalisis bagaimana sentimen masyarakat Indonesia pada mobil Wuling Almaz. Penggunaan Lexicon Based untuk membagi 3 klasifikasi yakni positif, netral, dan negatif. Hasil Lexicon Based pada 1000 komentar di akun Autonetmagz yakni 232 untuk positif, untuk netral berjumlah 456 dan negatif berjumlah 312. Dari kedua jurnal sebagai studi literatur tersebut didapati kedua jurnal tersebut hanya sampai kepada sentimen analisis dan tidak menjelaskan siapa aktor yang paling berpengaruh dalam komentar pada akun YouTube tersebut. Gap ini yang kemudian peneliti tertarik untuk menambahkan pada penelitian kali ini. Sehingga dengan penelitian ini, dapat menemukan kebaruan dalam proses analisis sentimen pada akun Youtube.

### B. Tempat & Waktu Penelitian

Penelitian ini dilaksanakan di berbagai lokasi yang mendukung produktivitas, seperti rumah untuk suasana yang kondusif, kampus dengan akses fasilitas akademik, perpustakaan untuk sumber referensi yang kaya, dan kafe yang menawarkan suasana inspiratif. Adapun waktu penelitian berlangsung sejak periode sebelum Ujian Tengah Semester (UTS) hingga berakhirnya masa Ujian Akhir Semester (UAS) di semester lima tahun ajaran 2024/2025, sesuai arahan dosen pengampu mata kuliah *Communication Network and Big Data*.

### C. Flowchart (Alur Model Pengembangan Sistem)

Berikut pada sub bab ini kami gambarkan alur model pengembangan dari penelitian kami yang memiliki 7 proses utama seperti pada gambar berikut:



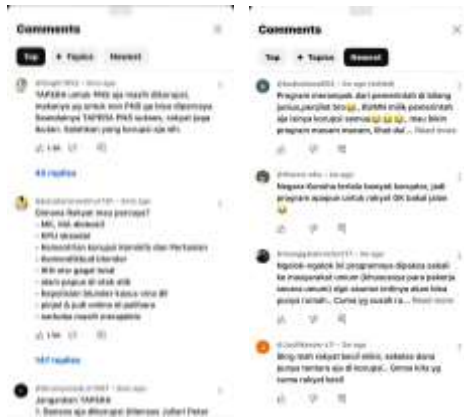
Gbr 1. Alur Model Pengembangan Sistem

Metode yang digunakan dalam penelitian ini diawali dengan proses *crawling data*, yakni pengambilan *raw data* komentar dari video *YouTube* Raymond Chin yang berjudul ‘Dukung TAPERA, Program JENIUS Pemerintah’ melalui *YouTube API* untuk mendapatkan informasi yang relevan secara *real-time*. Selanjutnya, data dimasukkan ke dalam proses *preprocessing*, yang mencakup beberapa langkah penting: Data yang telah terkumpul kemudian melewati tahap *cleaning* untuk menghapus data duplikat, simbol yang tidak diperlukan, serta elemen-elemen lain yang dapat mengganggu analisis, *data normalization* untuk menyeragamkan format teks, penghapusan *stopword* agar hanya menyisakan kata bermakna, *tokenization* untuk memecah teks menjadi unit analisis yang lebih kecil, dan *stemming* untuk mengembalikan kata ke bentuk dasarnya. Setelah itu, dilakukan pelabelan data (*labelling data*) menggunakan pendekatan *lexicon-based*, di mana setiap komentar diklasifikasikan berdasarkan polaritas sentimen—positif, negatif, ataupun netral.

Hasil dari proses ini akan divisualisasikan melalui diagram menggunakan *library* visualisasi seperti *Seaborn* dan *Matplotlib*, untuk memberikan pemahaman yang intuitif terhadap distribusi sentimen dan dinamika aktor di dalamnya. Lalu, peneliti juga akan menggunakan *WordCloud* untuk melihat berbagai kata yang sering muncul dalam komentar video youtube Raymond Chin yang berjudul *Dukung TAPERA, Program JENIUS Pemerintah*. Terakhir, akan dilakukan hasil analisa dari penelitian yang kami lakukan.

#### D. Dataset (Sumber Data)

Sumber data penelitian ini berupa komentar-komentar yang dikurasi dari video *YouTube* Raymond Chin yang secara spesifik membahas isu Tapera dengan judul ‘Dukung TAPERA, Program JENIUS Pemerintah’. Pemilihan dataset dilakukan secara teliti untuk menjamin keragaman opini, kekayaan perspektif, serta relevansi data dengan topik penelitian. Data ini mencakup 4 kolom penting yaitu teks komentar (*full\_text*), interaksi (*likes*), waktu unggahan (*created\_at*), dan *username* yang menjadi landasan untuk menganalisis dinamika sentimen publik serta identifikasi aktor yang memiliki pengaruh besar dalam diskusi pada kolom komentar video *Youtube* tersebut.



Gbr 2. Data Komentar Youtube Raymond Chin

Video yang di upload oleh *Raymond Chin* di *Youtube* pada 5 bulan lalu tepatnya pada 11 Juni 2024 memancing publik untuk mengeluarkan aspirasi mereka terhadap isu Tapera yang kian ramai dan menuai pro dan kontra. Dengan melakukan *crawling data*, dari 2105 komentar yang ada pada video tersebut terdeteksi 1809 komentar yang akan kami analisis lebih lanjut. Retang waktu yang kami gunakan untuk *crawling data* pada penelitian ini adalah komentar yang diposting sejak awal diposting sampai November 2024.

#### E. Alat dan Algoritma yang Digunakan

Dalam kerangka penelitian yang kami lakukan, ada berbagai *tools* dan algoritma pendukung yang membantu proses dari awal hingga akhir pemrosesan data guna memastikan keakuratan dalam analisis data, mulai dari alat pengumpulan data, alat pemrosesan dan pengembangan data serta algoritma analisis sentimen dan identifikasi aktor berpengaruh. Berikut adalah elaborasi dari setiap komponen yang kami gunakan dalam analisis:

##### 1. Youtube API

Dalam studi ini, *YouTube API* dioptimalkan sebagai instrumen utama untuk memperoleh data komentar pada video Raymond Chin secara spesifik video yang berjudul *Dukung TAPERA, Program JENIUS Pemerintah*, termasuk elemen metadata seperti waktu publikasi dan jumlah keterlibatan yang dihitung berdasarkan *likes*. Fitur unggulannya yang memungkinkan akses data berskala besar secara *real-time* menjadikan proses pengumpulan informasi lebih terorganisasi dan efisien. Dengan penerapan *API* ini, fondasi analisis sentimen dan identifikasi aktor berpengaruh dapat dibangun secara kokoh, sehingga dinamika interaksi digital tersaji dengan akurasi dan kedalaman analisis yang tinggi.

##### 2. Google Colab

Google Colab merupakan platform berbasis awan (*cloud-based*) yang mendukung peneliti dalam menjalankan skrip Python di lingkungan yang interaktif. Dengan akses ke unit pemrosesan grafis (*Graphics Processing Unit* atau GPU), *Google Colab* memungkinkan akselerasi pada proses pelatihan model, analisis data, dan visualisasi. Selain itu, fitur kolaboratifnya membuka ruang bagi peneliti untuk bekerja sama secara *real-time* dengan sesama tim peneliti, sehingga produktivitas dan efektivitas penelitian meningkat secara signifikan.

##### 3. Python

Bahasa pemrograman Python, tim peneliti pilih sebagai bahasa pemrograman utama dalam penelitian ini. Beragam pustaka (*libraries*) unggulan dimanfaatkan untuk mendukung seluruh tahapan pemrosesan dan analisis data secara optimal. *Pandas* dan *NumPy* digunakan untuk manipulasi serta pengelolaan data dalam skala besar dengan efisiensi tinggi, menjadikannya instrumen esensial dalam tahap *data preprocessing*. Selain itu, *Sastrawi*, pustaka yang dirancang khusus untuk bahasa Indonesia, berperan penting

dalam proses *stemming* dan pembersihan data sehingga memastikan kesiapan data untuk tahap analisis lanjutan. Dalam aspek visualisasi, *Seaborn* dan *Matplotlib* diimplementasikan untuk menghasilkan representasi visual yang tidak hanya informatif tetapi juga estetik, memungkinkan peneliti untuk menginterpretasikan pola dan distribusi data secara intuitif.

#### 4. *Lexicon-Based Algorithm*

Algoritma *lexicon-based* diterapkan dalam penelitian ini sebagai pendekatan sentimen yang memanfaatkan kamus sentimen preskriptif, mengandung leksikon polaritas positif, negatif, dan netral yang kami ambil melalui *github*. Algoritma ini menginterogasi teks hasil *preprocessing*—meliputi normalisasi, eliminasi *stopword*, tokenisasi, serta *stemming*—untuk mengekstrak terminologi yang relevan dengan bobot polaritas. Setiap entitas leksikal dianalisis melalui proses kalkulasi skor agregat, menghasilkan determinasi polaritas keseluruhan yang bersifat kuantitatif. Pendekatan ini unggul dalam memberikan interpretasi yang eksplisit tanpa kompleksitas komputasional seperti algoritma berbasis pembelajaran mesin, meskipun tetap menghadapi tantangan semantik, terutama dalam menangani ambiguitas atau konteks nonlinier yang melampaui cakupan leksikon.

#### F. *Metode yang Digunakan*

Penelitian ini memanfaatkan *YouTube API* untuk melakukan *crawling data* secara sistematis, memungkinkan pengumpulan komentar video Youtube Raymond Chin berjudul ‘Dukung TAPERA, Program JENIUS Pemerintah’. Analisis sentimen dilakukan menggunakan metode *Lexicon*, salah satu pendekatan dalam analisis teks yang berfokus pada penggunaan kosakata atau kamus kata (*lexicon*) untuk memahami atau menganalisis teks berdasarkan arti atau sentimen dari kata-kata yang terkandung di dalamnya. Selanjutnya, hal ini juga dibantu dengan kamus manual “*id opinion words*”, dimana kamus ini berisi kumpulan kata yang positif dan negatif yang berfungsi untuk memberikan preferensi kata-kata untuk opini bisa dikategorikan ke dalam sentimen positif atau negatif. Kombinasi metode serta kamus ini dirancang secara integratif untuk menghasilkan hasil yang lebih akurat serta memberikan hasil yang bisa menggambarkan untuk keperluan analisis komprehensif dalam berdinamika terkait opini publik di platform YouTube khususnya pada video yang dibuat oleh kreator Raymond Chin mengenai isu Tapera ini.

### III. HASIL DAN PEMBAHASAN

#### A. *Crawling Data*

##### 1. Fungsi *Crawling*

```
def fetch_comments(video_id):
    # Menginisialisasi objek
    replies = []

    url = f'https://www.youtube.com/watch?v={video_id}&comment=sort=oldest'

    # Memanggil API YouTube untuk mendapatkan komentar
    video_response = youtube.commentThread().list(part='snippet,videoInfo', videoId=video_id, maxResults=100)

    # Mendapatkan daftar komentar
    for item in video_response['items']:
        published = item['snippet']['topLevelComment']['snippet']['publishedAt']
        user = item['snippet']['topLevelComment']['snippet']['authorDisplayName']
        content = item['snippet']['topLevelComment']['snippet']['textDisplay']
        likeCount = item['snippet']['topLevelComment']['snippet']['likeCount']

        replies.append({'published': user, 'content': likeCount})

    # Menghitung jumlah komentar
    total_replies = len(replies)

    # Mengembalikan daftar komentar
    return replies

# Fungsi untuk memanggil API dan mendapatkan komentar
def fetch_comments(video_id):
    replies = fetch_comments(video_id)

    # Menghitung jumlah komentar
    total_replies = len(replies)

    # Mengembalikan daftar komentar
    return replies
```

Gbr 3. Fungsi *Crawling*

Proses dimulai dengan menginisialisasi layanan API menggunakan kunci autentikasi (`api_key`) dan memanggil fungsi `commentThreads().list` untuk mendapatkan komentar utama beserta balasannya. Data yang diambil meliputi waktu publikasi, nama pengguna, isi komentar, dan jumlah suka, yang kemudian disimpan dalam daftar `replies`. Jika komentar memiliki balasan, balasan tersebut juga diproses dan ditambahkan ke daftar. Kode ini juga menangani pengambilan komentar dalam beberapa halaman menggunakan `nextPageToken` untuk memastikan semua komentar diambil hingga halaman terakhir. Keluaran akhir adalah kumpulan semua komentar dan balasannya dalam bentuk terstruktur.

## 2. Proses Crawling



Gbr 4. Proses Crawling

Proses ini mengambil dan menyimpan komentar dari sebuah video pada platform tertentu. Kode tersebut memanfaatkan API untuk mengakses data komentar berdasarkan ID video yang spesifik. Hasilnya adalah daftar komentar yang berisi informasi seperti waktu komentar, nama pengguna, dan isi komentar.

### B. Data Preparation

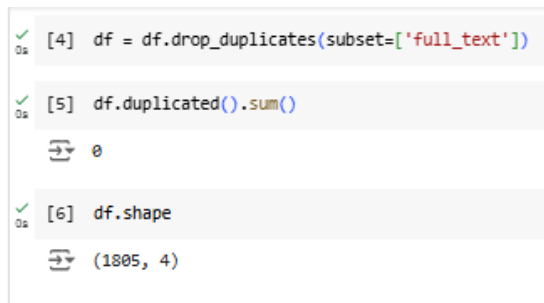


Gbr 5. Data Preparation

Proses membaca data dari file CSV menggunakan `pandas` dan melakukan eksplorasi awal data untuk memberikan ringkasan struktur `DataFrame`, termasuk jumlah baris, jumlah kolom, tipe data, serta nilai non-null di setiap kolom.

### C. Pre-processing Data

#### 1. Cleaning : Duplicate Removal



Gbr 6. Preprocessing: Cleaning Data



Normalisasi data teks berfungsi untuk menyamakan variasi penulisan kata-kata menjadi bentuk standar. Pengecekan dilakukan seperti mengganti kata-kata ke dalam bentuk yang baku, dan melihat apakah jenis data berbentuk string atau tidak.

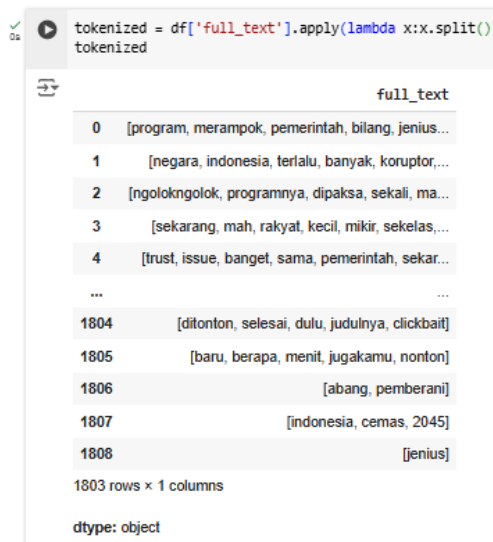
#### 4. Filtering : Stopwords Removal



Gbr 10. Proses Stopwords Removal

Mengolah data teks dalam sebuah DataFrame menggunakan pustaka Sastrawi. Hal ini bertujuan untuk membersihkan teks dengan menghapus kata-kata umum yang tidak memberikan informasi penting (*stopwords*).

#### 5. Filtering : Tokenize



Gbr 11. Proses Tokenisasi Data

Tokenisasi teks yang merupakan langkah awal dalam pengolahan data teks. Data teks mentah dalam kolom 'full\_text' dipecah menjadi daftar kata (token) menggunakan fungsi Python *split()* yang diterapkan pada setiap baris. Hasilnya adalah kolom baru berisi list kata-kata yang mewakili teks asli.

## 6. Filtering : Stemming



Gbr 12. Proses *Stemming Data*

Memproses teks untuk setiap kata menggunakan *stemming*, dan mengembalikan teks yang sudah di-stem dalam bentuk satu string menggunakan pustaka Sastrawi dan menyimpan hasilnya ke file CSV.

## D. Labelling Data



Gbr 13. Proses *Labelling Data*

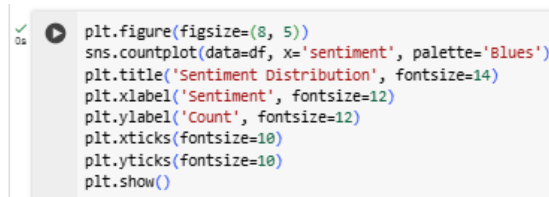


Gbr 14. Proses *Sentiment Analysis* dengan Metode *Lexicon*

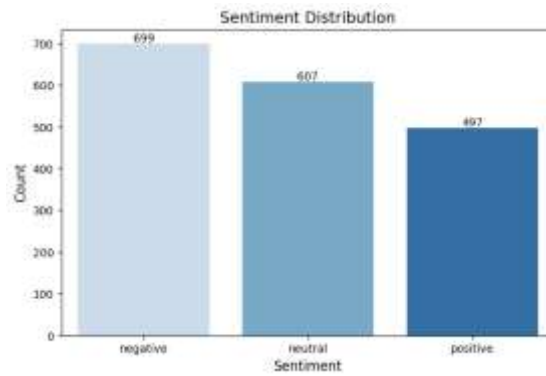
Proses analisis sentimen terhadap teks dalam sebuah DataFrame. Pendekatan ini menggunakan daftar kata positif dan negatif yang dibaca dari file eksternal untuk menilai polaritas (positif, negatif, atau netral) dari teks di kolom full\_text.

## E. Data Visualization

### 1. Sentiment Distribution



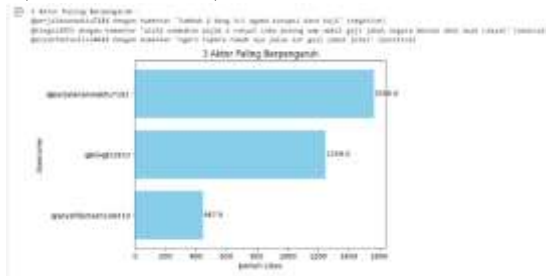
Gbr 15. Kode Python *Sentiment Distribution*



Gbr 16. Visualisasi *Sentiment Distribution*

Dilakukan visualisasi distribusi sentimen dari data menggunakan grafik batang (*bar chart*). Kode Python menggunakan *library* matplotlib dan seaborn untuk membuat grafik ini. Distribusi sentimen ditampilkan dalam tiga kategori yakni *negative* dengan jumlah 699, *neutral* dengan jumlah 607, dan *positive* dengan jumlah 497, dengan sumbu horizontal menunjukkan jenis sentimen dan sumbu vertikal menunjukkan jumlah (*count*) masing-masing kategori.

## 2. Actor Most Influential (based on likes)



Gbr 17. Visualisasi 3 *Most Influential Actor*

Gambar ini menampilkan grafik batang yang menunjukkan tiga pengguna paling berpengaruh berdasarkan jumlah *likes* yang diterima pada komentar mereka di media sosial. Pengguna @perjalananwaktu7181 dengan komentar berisi sentimen *negatif* mendapat jumlah *likes* tertinggi yakni 1566, diikuti oleh @kingki1953 dengan sentimen *neutral*, dan @arysfhantastico8419 dengan sentimen *positive*. Visualisasi ini membantu mengidentifikasi aktor kunci dalam diskusi serta bagaimana sentimen komentar mereka mempengaruhi keterlibatan audiens.

## 3. WordCloud

```
from wordcloud import WordCloud
text = "...-3424811 'full', 'best'...)
wordcloud = WordCloud(width=800, height=400, background_color='white').generate(text)
plt.figure(figsize=(10, 10))
plt.imshow(wordcloud, interpolation='bilinear')
plt.axis('off')
```

Gbr 19. Proses WordCloud



Gbr 20. Visualisasi Kata yang Sering Muncul

Visualisasi ini memanfaatkan *WordCloud* untuk merepresentasikan frekuensi leksikal dalam korpus teks secara grafis. Dari hasil analisa *WordCloud* ini dihasilkan beberapa kata dominan yang sering dipakai dalam komentar youtube Raymond Chin bertajuk ‘Dukung TAPERA, Program JENIUS Pemerintah’. Terlihat kata seperti ‘tapera’ dan ‘korupsi’ sering muncul, hal ini menandakan komentar yang ada sesuai terkait isu yang dibahas pada video tersebut. Lalu, kata lain seperti ‘program’ dan ‘perintah’ yang merupakan kata dasar dari pemerintah serta frasa-frasa lain yang juga sering muncul hadir untuk melengkapi pembahasan isu tersebut.

#### F. Hasil Analisis

Berdasarkan hasil analisis distribusi sentimen yang dilakukan pada data komentar terkait isu TAPERA pada komentar video youtube Raymond Chin yang berjudul *Dukung TAPERA, Program JENIUS Pemerintah*, terlihat bahwa opini masyarakat terbagi dengan cukup signifikan. Sentimen negatif mendominasi dengan jumlah 699 komentar, yang mencerminkan ketidakpuasan masyarakat terhadap program tersebut, diikuti oleh sentimen netral sebanyak 607 komentar dan sentimen positif yang lebih sedikit, yaitu 497 komentar yang mana setelah ditelusuri secara manual komentar positif lebih ditujukan kepada Raymond Chin sebagai konten kreator yang membuat video karena pembahasannya yang berbobot. Analisis lebih lanjut mengungkapkan bahwa komentar dengan sentimen negatif memperoleh keterlibatan yang tinggi, seperti yang terlihat pada komentar pengguna @perjalananwaktu7181 yang mendapatkan 1566 likes, menandakan tingginya perhatian publik terhadap ketidakpercayaan yang ada terhadap program ini. Fenomena ini mencerminkan adanya kegelisahan dan kecurigaan yang mendalam terkait kredibilitas dan efektivitas program TAPERA, yang banyak dipandang skeptis oleh masyarakat sebagai langkah dari pemerintah yang kurang transparan.

Isu utama yang berkembang dalam komentar masyarakat adalah terkait dengan kepercayaan publik terhadap pemerintah. Melalui visualisasi *WordCloud*, kata-kata seperti "tapera" dan "korupsi" muncul secara dominan, mengindikasikan bahwa banyak komentar yang mengaitkan program ini dengan ketidakpuasan terhadap praktik-praktik yang dianggap tidak transparan dan berpotensi merugikan masyarakat. Selain itu, kata-kata seperti "program" dan "pemerintah" juga sering muncul, yang menunjukkan bahwa diskusi seputar TAPERA tidak hanya berfokus pada program itu sendiri, tetapi juga menyentuh pada konteks pemerintahan yang dianggap kurang mampu meyakinkan rakyat akan niat baiknya. Sebagai hasilnya, wajar jika banyak komentar yang memiliki sentimen negatif, mencerminkan adanya *trust issue* yang mendalam di kalangan masyarakat terhadap kebijakan pemerintah terkait pengelolaan dana dan program yang dianggap tidak memenuhi ekspektasi publik.

#### IV. KESIMPULAN

Berdasarkan hasil keseluruhan langkah penelitian yang telah diambil maupun diuji dapat disimpulkan bahwa, di era digital era digital *Society 5.0*, inovasi teknologi telah menghadirkan berbagai solusi holistik untuk mengatasi tantangan sosial dan lingkungan, khususnya melalui

platform media sosial. Dengan pengambilan persepsi masyarakat secara komprehensif mengenai isu Tapera yang merugikan khalayak, Penulis juga menggunakan salah satu video yang dimana video ini berasal dari salah satu Podcaster yaitu Raymond Chin melalui platform Youtube sebagai dasar dalam pengambilan riset data mengenai sentimen publik yang berbentuk komentar video.

Penggunaan metode penelitian berupa Studi Literatur ini memfokuskan pada artikel jurnal yang kredibel serta berkaitan dengan Analisis Sentimen, *Lexicon-Based algorithm*, *Gap Analysis*, dan dinamika interaksi digital media sosial, khususnya platform Youtube. Pengambilan tempat & waktu penelitian ini dilakukan di berbagai lokasi yang mampu mendukung produktivitas sebagai contoh Rumah Komentar. Pengumpulan data set ini juga dilakukan secara menyeluruh dan teliti untuk menjamin keberagaman opini, kekayaan perspektif, serta kredibel maupun relevansi data dengan topik penelitian. Dalam melakukan penelitian, penulis menggunakan *Youtube API*, *Google Colab*, *Python*, *Lexicon-Based Algorithm* sebagai alat serta algoritma dalam mendukung keberlangsungan proses penelitian secara akurat dan objektif.

Metode yang disusun oleh penulis meliputi beberapa hal seperti, *Crawling Data*, *Data Preparation*, *Pre-processing Data*, *Labelling Data*, dan *Data Visualization* dalam memproses data serta menghasilkan hasil analisis melalui komentar video Youtube Raymond Chin. Melalui hasil analisis yang telah diproses secara sistematis dan struktural, dapat dilihat bahwa terdapat 699 komentar negatif secara dominan terkait ketidakpuasan masyarakat terhadap Tapera yang diikuti sebanyak 607 komentar netral dan 497 komentar positif yang terdapat di dalam video Raymond Chin. Dominasi sentimen negatif yang tinggi ini dapat diperlihatkan melalui komentar salah satu pengguna yaitu, @perjalananwaktu7181 dengan likes sebanyak 1566 di dalam isu Tapera ini.

## UCAPAN TERIMA KASIH

Segala puji dan syukur kami haturkan ke hadirat Tuhan Yang Maha Esa, atas limpahan rahmat, karunia, dan hikmah-Nya yang tiada terhingga, sehingga kami dapat menyelesaikan jurnal ini dengan judul "*Analisis Sentimen dan Aktor Berpengaruh di Komentar Video YouTube Raymond Chin Menggunakan Metode Lexicon*" sesuai dengan waktu yang telah direncanakan. Jurnal ini adalah wujud nyata dari dedikasi, integritas, dan usaha yang sungguh-sungguh dalam menjelajahi konsep-konsep mendalam yang menjadi inti dari mata kuliah *Communication Network and Big Data*.

Dengan segala kerendahan hati, kami menyampaikan penghormatan serta terima kasih yang sebesar-besarnya kepada berbagai pihak yang telah memberikan dukungan, arahan, dan inspirasi selama proses penyusunan jurnal ini. Ucapan terima kasih secara khusus kami tujukan kepada:

1. Tuhan Yang Maha Esa, yang senantiasa melimpahkan kemudahan dan keteguhan hati kepada kami dalam menghadapi setiap tantangan selama proses penyusunan jurnal ini.
2. Mas Putrawan Yuliandri, S.I.Kom., M.Si., selaku Dosen Pengampu Mata Kuliah *Communication Network and Big Data*, yang dengan penuh kesabaran dan keahlian memberikan panduan, bimbingan, serta wawasan yang sangat berharga bagi kami.
3. Rekan-rekan seperjuangan, yang telah menunjukkan solidaritas dan kontribusi berupa gagasan yang inovatif serta dukungan emosional yang menjadi penyemangat selama proses ini berlangsung.
4. Orang tua dan keluarga tercinta, atas doa yang tak pernah putus, kasih sayang yang tulus, dan dukungan moril maupun materil yang senantiasa menjadi pondasi kekuatan kami.
5. Seluruh pihak lain yang tidak dapat kami sebutkan satu per satu, yang dengan caranya masing-masing telah memberikan kontribusi berarti dalam penyelesaian jurnal ini.

Kami berharap karya ini mampu memberikan nilai tambah yang signifikan bagi dunia akademik maupun praktisi, khususnya dalam memahami dinamika analisis sentimen publik dan identifikasi aktor berpengaruh di media sosial. Kami juga sangat terbuka terhadap kritik, masukan, dan saran konstruktif yang dapat menjadi bahan refleksi untuk meningkatkan kualitas

karya ini di masa depan. Dengan rasa syukur dan optimisme, kami persembahkan jurnal ini sebagai salah satu sumbangsih kami terhadap pengembangan ilmu pengetahuan.

## DAFTAR PUSTAKA

- Afdhal, I., Kurniawan, R., Iskandar, I., Salambue, R., Budianita, E., & Syafria, F. (2022). Penerapan Algoritma Random Forest Untuk Analisis Sentimen Komentar Di YouTube Tentang Islamofobia. *Penerapan Algoritma Random Forest Untuk Analisis Sentimen Komentar Di YouTube Tentang Islamofobia*, 5(1), 122-130.
- Himawan, H., Putri, D. G., & Kaswidjanti, W. (2018). *METODE LEXICON BASED DAN SUPPORT VECTOR MACHINE UNTUK MENGANALISIS SENTIMEN PADA MEDIA SOSIAL SEBAGAI REKOMENDASI OLEH - OLEH FAVORIT*, 1, 235–244.
- Nasrullah, R. (2015). *Media Sosial: Perspektif Komunikasi, Budaya dan Sosioteknologi*. Bandung: Simbiosis Rekatama Media.
- Prasetya, H., Buntoro, G. A., & Mustikasari, D. (2020). Analisis Sentimen Pada Channel Autonetmagz Terhadap Review Mobil Almaz 2019 Dengan Metode Naive Bayes Classifier Dan Lexicon Based. *KOMPUTEK*, 4(1), 58-73.
- Rahmadi, R., & Rajagede, R. A. (2021). Analisis Sentimen Politik Berdasarkan Big Data dari Media Sosial Youtube: Sebuah Tinjauan Literatur. *Automata*, 2(1).
- Sanjaya, G., & Lhaksana, K. M. (2020). Analisis Sentimen Komentar YouTube tentang Terpilihnya Menteri Kabinet Indonesia Maju Menggunakan Lexicon Based. *eProceedings of Engineering*, 7(3).
- Thomas, S., Yuliana, & Noviyanti, P. (2021). Study analisis metode analisis sentimen Pada YouTube. *Journal of Information Technology*, 1(1), 1–7. <https://doi.org/10.46229/jifotech.v1i1.20>
- Yudianto, M. R. A., Rahim, A., Sukmasetya, P., & Hasani, R. A. (2022). *Perbandingan Metode Support Vector Machine Dengan Metode Lexicon Dalam Analisis Sentimen Bahasa Indonesia*, 6, 7–13. <https://doi.org/https://doi.org/10.36294/jurti.v6i1.2537>
- Yuliana, Y., & Noviyanti, P. Study Analisis Metode Analisis Sentimen pada YouTube. *Journal of Information Technology*, 1(1), 1-7. <https://github.com/masdevid/ID-OpinionWords/tree/master> diakses pada Selasa, 10 Desember 2024 pukul 19:25
- <https://wearesocial.com/id/blog/2024/01/digital-2024/> diakses pada Minggu, 8 Desember 2024 pukul 12:27
- Huwaida, S. F., Kusumawati, R., & Isnaini, B. (2024). Analisis sentimen komentar youtube terhadap pemindahan ibu kota negara menggunakan metode Naïve Bayes. *Jambura Journal of Informatics*, 6(1), 26-39. <https://doi.org/10.37905/jji.v6i1.24718>
- Amrullah, F., & Solichin, A. (2024). Analisis Emosi Pada Live Chat Youtube'Mata Najwa: 3 Bacapres Bicara Gagasan'Menggunakan Pendekatan Lexicon dan Algoritma Naive Bayes. *Jurnal Ticom: Technology of Information and Communication*, 12(3), 121-128. <https://doi.org/10.70309/ticom.v12i3.132>
- Syah, R. I., Hoiriyah, H., & Walid, M. (2023). Analisis Sentimen Pengguna Media Sosial Terhadap Aplikasi M-Health Peduli Lindungi Dengan Metode Lexicon Based Dan Naïve Bayes. *Indonesian Journal of Business Intelligence (IJUBI)*, 6(1). <https://dx.doi.org/10.21927/ijubi.v6i1.3275>