

ANALISIS SENTIMEN TIKTOK: WAJIB MILITER DENGAN METODE LEXICON BASED DAN NAIVE BAYES CLASSIFIER

Arpan Mualief Saprizal✉, Nor Anisa

Fakultas Sains & Teknologi, Universitas Sari Mulia, Banjarmasin, Indonesia

Email: arpan.mualief@student.unism.ac.id

ABSTRACT

The issue of conscription in Indonesia has sparked a heated debate among the public, especially on the social media platform TikTok. This study aims to analyze public sentiment on the issue through analysis of TikTok user comments. The method used is lexicon-based sentiment analysis. Data of 5,212 comments were collected using web scraping techniques with the keyword "conscription in Indonesia". The results of the analysis showed that the majority of comments (53.28%) were positive, followed by neutral comments (35.79%), and negative comments (10.92%). This finding indicates that there is considerable support for the issue of military service among TikTok users. The research process includes data collection, data processing, sentiment analysis using a lexicon-based approach, and visualization of results. The results of this study are expected to provide a clearer picture of public perception of the issue of military conscription in Indonesia.

Keyword: *Military Conscription, Sentiment-Analysis, Lexion-Based, Naive Bayes Classifier, Wordcloud.*

ABSTRAK

Isu wajib militer di Indonesia telah memicu perdebatan hangat di kalangan masyarakat, terutama di platform media sosial TikTok. Penelitian ini bertujuan untuk menganalisis sentimen publik terhadap isu tersebut melalui analisis terhadap komentar pengguna TikTok. Metode yang digunakan adalah lexicon-based sentiment analysis. Data sebanyak 5.212 komentar dikumpulkan menggunakan teknik web scraping dengan kata kunci "wajib militer di Indonesia". Hasil analisis menunjukkan bahwa mayoritas komentar (53,28%) bernada positif, diikuti oleh komentar netral (35,79%), dan negatif (10,92%). Temuan ini mengindikasikan adanya dukungan yang cukup besar terhadap isu wajib militer di kalangan pengguna TikTok. Proses penelitian meliputi pengumpulan data, pengolahan data, analisis sentimen menggunakan lexicon--based approach, dan visualisasi hasil. Hasil penelitian ini diharapkan dapat memberikan gambaran yang lebih jelas mengenai persepsi publik terhadap isu wajib militer di Indonesia.

Kata Kunci: *Wajib Militer, Analisis-Sentimen, Lexion-Based, Naive Bayes Classifier, Wordcloud.*

PENDAHULUAN

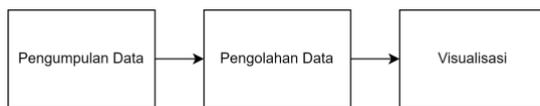
Sosmed atau sosial media telah menjadi platform utama bagi masyarakat untuk mengekspresikan opini dan pandangan mereka terhadap berbagai isu, termasuk isu sosial dan politik Faoziah, Cindytiast Putri. TikTok, sebagai salah satu platform media sosial paling banyak digunakan, memungkinkan jutaan pengguna untuk membuat dan membagikan video pendek dengan mudah (Rachmawaty, 2021). termasuk yang berkaitan dengan isu-isu terkini(Safitri, 2024). Salah satu isu yang sering diperdebatkan di Indonesia adalah mengenai wajib militer. Munculnya isu ini telah memicu perbincangan yang hangat di berbagai kalangan, termasuk di kalangan pengguna TikTok. Analisis sentimen, atau yang sering disebut penambangan opini, adalah proses mengidentifikasi, mengekstrak, dan mengklasifikasikan pendapat atau sikap yang tersirat dalam suatu teks(Arief & Imanuel, 2019). Analisis

sentimen memungkinkan kita untuk melacak secara sistematis bagaimana masyarakat merespon kebijakan publik dengan menganalisis sentimen yang terungkap dalam unggahan media sosial. Analisis sentimen adalah proses mengungkap sikap atau opini yang tersembunyi dalam teks, baik itu positif, negatif, maupun netral (Sanjaya et al., 2024). Sikap-sikap ini umumnya diungkapkan oleh pengguna internet di media sosial sebagai bentuk ekspresi pribadi terhadap suatu topik atau produk (Saber & Saad, 2017). Melalui pendekatan leksikon, penelitian ini berupaya mengklasifikasi opini publik yang tertuang dalam komentar masyarakat terkait isu wajib militer ke dalam kategori sentimen positif, negatif, atau netral.

METODOLOGI PENELITIAN

Penelitian ini menganalisis sentimen masyarakat terhadap isu wajib militer menggunakan metode leksikon. Metode leksikon adalah suatu

pendekatan dalam analisis sentimen yang menggunakan kamus atau daftar kata yang telah ditentukan sebelumnya untuk mengidentifikasi sentimen (positif, negatif, atau netral) dalam sebuah teks (Mukti et al., 2023). Sentimen ini dikategorikan menjadi positif, negatif, dan netral. Proses penelitian meliputi pengumpulan data, analisis sentimen, dan visualisasi (lihat Gambar 1).



Gambar 1. Tahap-Penelitian

Pengumpulan Data

Data yang di gunakan merupakan komentar-komentar masyarakat dari video tiktok dengan kata kunci Wajib Militer di Indonesia sebanyak 5.212 data pada tanggal 11 November 2024. Proses scraping data dilakukan dengan Web Scraping. List data video TikTok kemudian diubah menjadi format tabel (data frame) untuk memudahkan pengolahan data. Dari tabel tersebut, kami mengekstrak teks komentar dari kolom 'F5' untuk analisis lebih lanjut. List Video Tiktok merupakan kumpulan data yang berisi informasi tentang isi video TikTok, seperti deskripsi video, jumlah views, jumlah likes, musik yang digunakan, dll(Zhao, 2022).

Pengolahan

Tujuan utama tahap pengolahan data adalah menyiapkan data komentar TikTok untuk analisis sentimen. Proses ini meliputi pemilihan kolom komentar, normalisasi teks (mengubah huruf menjadi huruf kecil, menghapus tanda baca, angka, dan kata-kata tidak penting), serta pembersihan data dari noise seperti URL, mention, dan emoticon.

pengolahan information adalah proses yang sangat penting dalam period informasi saat ini. Dengan mengolah information dengan baik, kita dapat memanfaatkan informasi yang ada untuk meningkatkan kualitas hidup dan membuat dunia menjadi lebih baik (Cholik, 2021).

Setelah selesai dilanjutkan dengan analisis sentimen untuk pengklasifikasian data komentar tiktok. Analisis sentimen dalam penelitian ini menggabungkan metode klasifikasi Naïve Bayes dengan pendekatan berbasis leksikon. Kombinasi ini memungkinkan klasifikasi sentimen secara efektif (Daniati & Utama, 2023). Metode ini menganalisis ulasan dengan cara menghitung kemunculan kata-kata yang memiliki muatan emosional dan mengkombinasikannya dengan informasi nilai rating

untuk menentukan sentimen keseluruhan (Fathullah et al., 2020).

Visualisasi

Untuk menyajikan hasil analisis sentimen secara lebih menarik dan intuitif, tahap akhir penelitian ini melibatkan visualisasi data menggunakan diagram histogram dan word cloud. Word cloud memungkinkan kita untuk mengidentifikasi kata-kata kunci yang paling sering muncul dalam teks, dengan ukuran font yang mencerminkan frekuensi kemunculannya (Charisma et al., 2024).

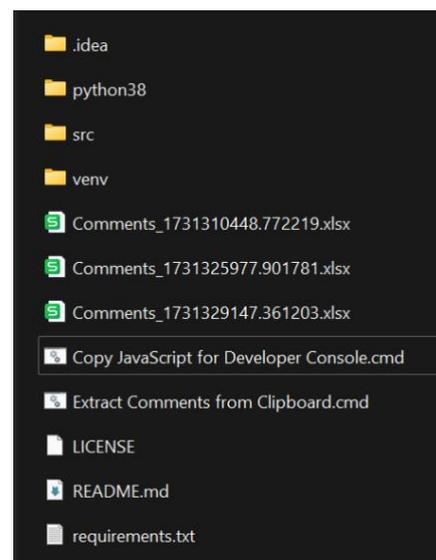
HASIL DAN PEMBAHASAN

Hasil Pengumpulan Data

Pengambil data dengan Web Scraping pada tanggal 11 November 2024 dengan kata kunci Wajib Militer di Indonesia telah mengumpulkan sebanyak 5.212 buah komentar video tiktok. Web Scraping di lakukan dengan memberikan kode di Console Tiktok yang dimana codenya berfungsi menyalin list data tiktok ke dalam file .xlsx

Langkah-langkah Web Scraping.

1. membuka console di bagian komentar tiktok,
2. di dalam console di lakukan pekodean.
3. Code ini berfungsi menyalin semua isi data pada video ke dalam clipboard.
4. Setelah itu extract salinan clipboard menggunakan cmd dan hasil extract akan mengoutput file .xlsx seperti gambar 2.



Gambar 2. Web Scraping

Berikut gambar data setelah melakukan web scraping.

Gambar 3. Data Hasil Web Scraping

Selanjutnya hanya kolom F5 yang akan di ubah menjadi data frame agar dapat diproses pada analisis selanjutnya.

Hasil Pengolahan

Proses terdiri antara lain:

1. Pengolahan kolom F5 ke Data Frame
2. Mengubah semua huruf menjadi huruf kecil (Case Folding)
3. Menstemmer Komentar
4. Pembersihan Teks dengan Menggunakan code.

```
def clean_text(text):
    text = re.sub(r"http\S+|www\S+|https\S+", "", text) # Hapus URL
    text = re.sub(r"[^a-zA-Z\s]", "", text) # Hapus tanda baca
    text = re.sub(r"(\s){2,}", r" ", text) # Hapus huruf berulang
```

Gambar 4. Code Cleaning Text

Pembersihan teks (*content cleaning*) adalah proses merapikan dan menyederhanakan teks mentah agar lebih mudah dianalisis. Proses ini melibatkan penghapusan atau pengubahan bagian-bagian teks yang tidak relevan atau mengganggu, seperti tanda baca yang tidak perlu, kata-kata halt, atau kesalahan pengetikan (NA'IM, 2023).

5. Kombinasi Normalisasi dan Kamus Singkatan.

Tabel 1. Data Frame dari kolom F5

Kamus Singkatan

“ngab” : ”bang”,
 “gua” : ”saya”,
 “wamil” : ”wajib militer”,
 “kek” : ”seperti”,

6. Memberikan Custom lexion

Selanjutnya akan ditampilkan dalam bentuk tabel yang isi datanya dari nomor yang menampak perubahan saja, dikarenakan tidak semua komentar membutuhkan processing.

Setelah kolom F5 di ubah menjadi Data Frame, selanjutnya dilakukan Case Folding untuk mengubah semua huruf menjadi huruf kecil (Lestandy et al., 2021) seperti pada tabel 2.

Tabel 2. Case Folding

Sebelum	Sesudah
1. Berguna ngab minimal bagian antar" makanan aja	1. berguna ngab minimal bagian antar" makan aja
2. ??Setuju.	2. ??setuju.
3. Gua menolak keras wajib militer	3. gua menolak wajib militer
4. tubuhku kecil, tinggi 156cm berat 36kg. orang kek aku apa berguna dimedan perang?	4. tubuh kecil, tinggi 156cm berat 36kg. orang kek aku apa berguna medan perang?

Stemming teks adalah proses mereduksi kata-kata ke bentuk dasarnya (root word) dengan menghilangkan akhiran seperti imbuhan atau infleksi. Tujuannya adalah untuk mengelompokkan kata-kata yang memiliki makna serupa menjadi satu kelompok, sehingga memudahkan dalam analisis teks.

Tabel 3. Stemmer Text

Sebelum	Sesudah
1. berguna ngab minimal bagian antar" makanan aja	1. guna ngab minimal bagian antar" makan aja
2. ??setuju.	2. ??tju.
3. gua menolak keras wajib militer	3. gua tolak wajib militer
4. tubuhku kecil, tinggi 156cm berat 36kg. orang kek aku apa berguna dimedan perang?	4. tubuh kecil, tinggi 156cm berat 36kg. orang kek aku apa guna medan perang?

Tabel 4. Pembersihan Text

Sebelum	Sesudah
1. guna ngab minimal bagian antar" makan aja	1. guna ngab minimal bagian antar makan aja
2. ??tju	2. tju
3. gua tolak wajib militer	3. gua tolak wajib militer
4. tubuh kecil, tinggi 156cm berat 36kg. orang kek aku apa guna medan perang?	4. tubuh kecil tinggi cm berat kg orang kek aku apa guna medan perang

Normalisasi adalah proses mengubah information ke dalam bentuk standar atau 'normal' agar lebih mudah diolah dan dianalisis. Dalam konteks pemrosesan bahasa alami, normalisasi sering dilakukan pada teks untuk menyederhanakan

- embedding pada twitter. *Journal of Information System Management (JOISM)*, 4(2), 125–131.
- Fathullah, N. S., Sari, Y. A., & Adikara, P. P. (2020). Analisis Sentimen Terhadap Rating dan Ulasan Film dengan menggunakan Metode Klasifikasi Naïve Bayes dengan Fitur Lexicon-Based. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 4(2), 590–593.
- Lestandy, M., Abdurrahim, A., & Syafa'ah, L. (2021). Analisis Sentimen Tweet Vaksin COVID-19 Menggunakan Recurrent Neural Network dan Naïve Bayes. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(4), 802–808.
- Mukti, A., Hadiyanti, A. D., Nurlaela, A., & Panjaitan, J. (2023). Sistem Analisa Sentiment Bakal Calon Presiden 2024 Menggunakan Metode NLP Berbasis Web. *SOSCIED*, 6(1), 128–140.
- NA'IM, M. (2023). *Analisis sentimen kinerja kepemimpinan bupati dari data komentar menggunakan metode naïve bayes classifier (Studi: Kecamatan Laren Kabupaten Lamongan 2020-2024)*. Universitas Islam Lamongan.
- Rachmawaty, A. (2021). Optimasi media sosial dalam meningkatkan penjualan di masa pembatasan sosial berskala besar. *TEMATIK*, 8(1), 29–44.
- Saberi, B., & Saad, S. (2017). Sentiment analysis or opinion mining: A review. *Int. J. Adv. Sci. Eng. Inf. Technol*, 7(5), 1660–1666.
- Safitri, R. D. (2024). Peran Tim Media Center PKB dalam Pemenangan Pemilu 2024 (Studi Kasus di DPC PKB Jember). *FORMULA Jurnal Administrasi Publik*, 1(1), 79–92.
- Sanjaya, R., Tohidi, E., Wahyudi, E., & Kaslani, K. (2024). Analisis Sentimen Terhadap Berhentinya Tiktokshop Pada Media Sosial Twitter Menggunakan Algoritma Naïve Bayes. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(1), 507–514.
- Zhao, B. (2022). Web scraping. In *Encyclopedia of big data* (pp. 951–953). Springer.