



Analisis Sentimen Komentar Video Putusan MA Terkait Kaesang Menggunakan Metode Naive Bayes

Abim Febri Hananto¹, Raihan Canggih Panilih², Reihan Setya Banda Syah Putra³,
Tariq Tariq⁴, Wildan Setiawan⁵

¹⁻⁵Universitas Duta Bangsa Surakarta, Indonesia

Alamat: Fakultas Ilmu Komputer Universitas Duta Bangsa Surakarta, Jl. Bhayangkara No.55, Tipes,
Kec. Serengan, Kota Surakarta, Jawa Tengah 57154

Email: 220101002@mhs.udb.ac.id^{1*}, 220101033@mhs.udb.ac.id², 220101034@mhs.udb.ac.id³,
220101037@mhs.udb.ac.id⁴, 220101040@mhs.udb.ac.id⁵

Abstract. Political dynasty is a political power exercised by a group of people who are related by family, with the aim of obtaining power and ensuring that this power remains within the group by passing it on to other family members. This study conducts a sentiment analysis on comments related to the Supreme Court decision which is believed to pave the way for Kaesang Pangarep in support of Jokowi's political dynasty. Sentiment analysis is carried out using the Naive Bayes method, a commonly used algorithm for text classification based on probability. The data used consists of comments from videos taken from social media platforms. These comments are then categorized into positive, negative, and neutral sentiments. The results of the study show the distribution of public sentiment towards this issue, providing an overview of how the public responds to the decision. The Naive Bayes method is chosen for its simplicity and its ability to provide reasonably accurate results in text analysis.

Keywords: Sentiment analysis, Naive Bayes, Political dynasty, Supreme Court, Kaesang Pangarep.

Abstrak. Politik dinasti merupakan kekuasaan politik yang dijalankan oleh sekelompok orang yang masih terkait dalam hubungan keluarga, dengan tujuan untuk memperoleh kekuasaan dan menjaga agar kekuasaan tersebut tetap berada di pihak kelompok tersebut. Pada penelitian ini, dilakukan analisis sentimen terhadap komentar video terkait putusan Mahkamah Agung yang dianggap memuluskan jalan Kaesang Pangarep demi dinasti politik Jokowi. Analisis sentimen dilakukan menggunakan metode Naive Bayes, sebuah algoritma yang umum digunakan untuk klasifikasi teks berdasarkan probabilitas. Data yang digunakan berupa komentar dari video yang diambil dari platform media sosial. Komentar-komentar ini kemudian dikategorikan menjadi sentimen positif, negatif, dan netral. Hasil penelitian menunjukkan distribusi sentimen masyarakat terhadap isu tersebut, yang dapat memberikan gambaran tentang bagaimana publik merespon keputusan ini. Metode Naive Bayes dipilih karena kesederhanaannya dan kemampuannya untuk memberikan hasil yang cukup akurat dalam analisis teks.

Kata kunci: Analisis sentimen, Naive Bayes, Politik dinasti, Mahkamah Agung, Kaesang Pangarep.

1. LATAR BELAKANG

Media sosial sangat penting di era internet saat ini untuk membentuk opini publik dan menyebarkan informasi dengan cepat. Media sosial, yang muncul bersamaan dengan perkembangan teknologi web 2.0, telah berkembang menjadi platform yang memungkinkan orang berkomunikasi dan berpartisipasi aktif dalam pembuatan dan pertukaran konten. YouTube adalah salah satu platform yang paling terkenal di antaranya menjadi lebih dari sekadar hiburan, menjadi sarana untuk mengajar dan berbagi informasi[1]. Menurut laman berita solopos.com, ada 170 juta pengguna, atau 93,8 persen dari 181,9 juta pengguna internet berusia 16-64 tahun. Di sisi lain, orang yang menggunakan situs YouTube juga dapat memberikan komentar, like, dislike, dan komentar pada video yang ditayangkan di

media youtube.[2]telah berkembang menjadi tempat diskusi yang dinamis tentang berbagai topik, termasuk politik. Putusan Mahkamah Agung Muluskan Jalan Kaesang Demi Dinasti Jokowi adalah salah satu perselisihan yang paling hangat. Dengan rencana pencalonannya ini, Kaesang, seorang pengusaha muda dan putra bungsu Presiden Joko Widodo, berhasil menarik perhatian publik dan media.

Pencalonan Kaesang tidak hanya memicu diskusi di media massa, tetapi juga menjadi topik hangat di Youtube karena platform ini memungkinkan pengguna untuk menyampaikan pendapat mereka secara langsung dan secara real-time, memberikan gambaran sentimen publik terhadap isu tersebut. Tujuan analisis sentimen Youtube adalah untuk mengidentifikasi dan memahami bagaimana masyarakat bereaksi terhadap berita tentang putusan Mahkamah Agung dan pencalonan Kaesang, apakah itu dengan cara yang positif atau negatif.

Kami akan mempelajari tanggapan dan perasaan pengguna Youtube terhadap putusan Mahkamah Agung terhadap pencalonan Kaesang sebagai Wakil Gubernur Jakarta. Dengan menggunakan metode analisis sentimen, kita dapat mendapatkan pemahaman yang lebih mendalam tentang pandangan publik, menemukan tema-tema utama yang dibahas, dan mengevaluasi pendapat dan argumen yang muncul di masyarakat. Analisis ini tidak hanya memberikan gambaran tentang persepsi publik, tetapi juga dapat membantu pengambil kebijakan dan tim kampanye memahami dan menanggapi aspirasi masyarakat dengan lebih baik.

Selain itu, artikel ini akan membahas hasil analisis sentimen terhadap dinamika politik di Jakarta, serta bagaimana sentimen di media sosial dapat memengaruhi keputusan politik dan strategi kampanye. Analisis sentimen adalah metode untuk melihat respons komentar masyarakat di media YouTube. Salah satu bagian dari pengolahan teks adalah analisis sentimen, yang secara otomatis mengkategorikan tulisan atau opini ke dalam kategori positif, netral, atau negatif[3]. Dengan demikian, tujuan dari artikel ini adalah untuk mendapatkan apa tanggapan masyarakat tentang bagaimana masalah pencalonan Kaesang Pangarep sebagai Wakil Gubernur Jakarta khususnya di platform Youtube

2. KAJIAN TEORITIS

2.1 Youtube

YouTube adalah platform media sosial yang menampilkan video dengan durasi tertentu yang dibuat oleh seniman video (kreator YouTube) dari berbagai genre. Karena tingginya antusias masyarakat dalam menggunakan YouTube, media sosial ini semakin banyak digunakan oleh berbagai kalangan. Youtube merupakan salah satu media sosial terpopuler di

Indonesia. Berdasarkan situs berita Solopos.com, memiliki 170 juta pengguna atau mewakili 93,84,444 persen dari total populasi pengguna Internet berusia 16 hingga 64 tahun sebanyak 181,9 juta orang. Di sisi lain, halaman YouTube juga memungkinkan Anda memberikan feedback berupa suka, tidak suka, dan komentar untuk video yang ditayangkan di media YouTube. Izinkan pemirsa berinteraksi dengan video Anda dan mengungkapkan pendapat mereka. Tentunya berdasarkan reaksi penonton, komentar tersebut bisa saja berisi komentar yang menyemangati dan memotivasi pembuatnya, dan bisa juga dalam bentuk SARA (Suku, Ras, Agama, Antar golongan) untuk YouTube komentar pada. Pembuatnya sendiri YouTube, salah satu platform streaming video terpopuler di dunia, menampilkan berbagai macam video hiburan dan pendidikan.[4]

2.2 Analisis Sentimen

Penambangan opini atau analisis sentimen adalah bidang studi di mana orang melakukan analisis sentimen komentar YouTube tentang ceramah pendapat mereka tentang hal-hal seperti barang dan jasa, organisasi, individu, masalah, peristiwa, dan topik. Karena sebagian besar penelitian berfokus pada pendapat seseorang yang mengekspresikan atau menyiratkan sentimen positif atau negatif, analisis sentimen sekarang menjadi pusat penelitian media sosial. Akibatnya, penelitian yang berkaitan dengan analisis sentimen mungkin memiliki pengaruh yang signifikan terhadap proses pengolahan bahasa alami (NLP). Pendapat publik telah menjadi penting dalam pemasaran, hubungan masyarakat, dan perusahaan kampanye politik selama bertahun-tahun. Pengaruh dan keuntungan analisis sentimen telah mendorong lebih banyak penelitian dan pemanfaatannya. Bahkan di Amerika Serikat, ada sekitar 20 hingga 30 bisnis yang fokus pada layanan analisis sentimen.[5].

2.3 Naive Bayes

Pembelajaran mesin Naive Bayes menggunakan probabilitas. Membuat asumsi yang sangat kuat (naif) tentang seberapa independen setiap kondisi atau peristiwa adalah ciri utama pengklasifikasi Naive Bayes. Ini adalah algoritma pembelajaran yang populer untuk klasifikasi teks dan pembelajaran mesin. Metode ini melibatkan perhitungan probabilitas. Pada penelitian ini, klasifikasi dokumen dilakukan dalam dua tahap. Dokumen dengan kategori yang sudah diketahui dilatih pada tahap pertama, dan dokumen dengan kategori yang belum diketahui diuji pada tahap kedua. Pengklasifikasi Naive Bayes sering digunakan karena kualitasnya yang luar biasa, meskipun sederhana. Dengan menggunakan teknik Bayes sederhana ini, orang dapat menemukan probabilitas Bayesian dan membuat pernyataan probabilitas.

Naive Bayes dapat dirumuskan sebagai Persamaan

$$P(H|X) = \frac{P(H) \prod P(X_i|H)}{P(X)}$$

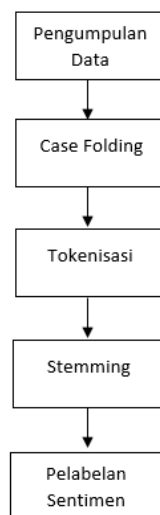
Naive Bayes Algorithm

Gambar 2.1 Persamaan

Sistem klasifikasi Naive Bayes menggunakan model klasifikasi probabilistik. Selain itu, ini adalah algoritma yang memiliki kemampuan untuk memperjelas data. Metodologi Penelitian: Beberapa modul utama digunakan dalam penelitian ini: pengumpulan data, pengolahan, klasifikasi, dan evaluasi.

3. METODE PENELITIAN

Dalam penelitian ini, penulis menelusuri opini masyarakat di media sosial Youtube terhadap Dejavu Pak Kesang sebagai calon wakil gubernur Jakarta tahun 2024. Data yang digunakan berupa komentar youtube. Dalam penelitian ini, kami melakukan beberapa langkah metodologis untuk memperoleh dan menganalisis data menggunakan alat.



Gambar 3.2 Alur Penelitian

3.1 Pengumpulan Data:

Pada langkah awal penelitian, para peneliti mengidentifikasi komentar untuk kumpulan data. Dalam penelitian ini, data berasal dari platform YouTube. Komentar yang digunakan

dalam penelitian ini adalah Komentar dari channel Youtube CNN Indonesia. Para peneliti mengumpulkan data menggunakan metode pengumpulan menggunakan web Scrapping[6].

3.2 Case Folding

Case Folding merupakan salah satu jenis metode preprocessing teks yang bertujuan untuk mengubah semua huruf dalam dokumen menjadi huruf kecil.[7]. Selain itu, Supaya data teks yang diproses dalam bentuk yang sama, case folding bertujuan untuk mengembalikan setiap kata ke dalam bentuk huruf kecil. Pembersihan teks adalah proses pembersihan kata yang tidak diinginkan untuk mempermudah proses klasifikasi. Karakter dan kata yang dihilangkan [8].

3.3Cleaning

Cleaning merupakan tahapan yang bertujuan menghilangkan karakter atau symbol link url (<http://situs.com>), username atau mention(@username), hastage(#), dan emoticon [9]

3.4 Tokenizing

Tokenizing merupakan salah satu jenis metode preprocessing teks. Tokenizing melibatkan pembagian teks menjadi token individual (biasanya kata atau tanda baca). Tokenizing adalah langkah pra-pemrosesan yang umum ketika bekerja dengan data teks, karena memfasilitasi pemrosesan dan analisis teks.

3.5 Stemming

Stemming adalah tindakan untuk menggabungkan berbagai variasi morfologis dalam kata dan kalimat menjadi satu bentuk dasar yang sama[10]. Stemming adalah dasar dari suatu teknik pemrosesan bahasa alami untuk mendapatkan informasi kembali (Information Retrieval) yang efisien dan umum digunakan[11]

3.6 Pelabelan

Labeling yang berasal berasal dari kata "label", yang berarti himpunan karakter atau karakter yang digunakan untuk mengidentifikasi variabel atau bagian tertentu dari data atau berkas.

4. HASIL DAN PEMBAHASAN

4.1 Pengumpulan Data

Dalam penelitian ini diperoleh dari komentar pengguna YouTube dalam tayangan Dejavu Putusan Mahkamah Agung Muluskan Jalan Kaesang Demi Dinasti Jokowi. Data tersebut diambil dengan cara scrapping menggunakan Bahasa pemrograman python. Dalam penelitian ini diambil data sebanyak 1186 komentar. Hasil scrapping dipaparkan di Gambar 3.

Username	Comment	Sentiment
@sarmilih1233	biasa itu lagi ada kuasa galak ntar turun jadi tikus	negatif
@tamamesta	pdi pasang anies-ahok aja pasti menang	negatif
@christinesupami2365	presiden mulut mencla mencle terulang lagi mengerikan karena mampu nyogok ma dan mk mulutnya mencla mencle dengan cara merubah uu dasar	negatif
@adeirawan4631	jakarta itu kota kecil jadi kaesang anak muda pasti bisa dia kan udah banyak pengalaman	positif
@fitriyanurulh4967	ditunggu2	netral

Gambar 4.3 Data Comment

4.2 Case Folding

Selanjutnya adalah tahap case folding dimana tahap ini juga termasuk tahap text processing dan menggunakan google collab dengan bahasa pemrograman python. Tahap ini mengubah huruf besar menjadi huruf kecil.

```
[50] df['comment'] = df['comment'].str.lower()
```

Gambar 4.4 Kodingan Case Folding

4.3 Cleaning

Selanjutnya tahap cleaning ini juga termasuk tahap text processing. dimana tahap ini dilakukan untuk pembersihan kata

```
[51] df = df.drop_duplicates(subset=['username', 'comment'])
df = df.dropna()

def clean_youtube_text(text):
    if isinstance(text, str): # Check if text is a string
        text = re.sub(r'@[a-zA-Z0-9_]+', '', text)
        text = re.sub(r'#[a-zA-Z0-9_]+', '', text)
        text = re.sub(r'http://[a-zA-Z0-9_]+', '', text)
        text = re.sub(r'https://[a-zA-Z0-9_]+', '', text)
        text = re.sub(r'[^a-zA-Z0-9 ]+', '', text)
        text = re.sub(r'(\s+)', ' ', text).strip()
        return text
    else:
        return "" # Return an empty string for non-string values

df['comment'] = df['comment'].apply(clean_youtube_text)
```

Gambar 4.5 Kodingan untuk Cleaning

4.4 Tokenizing

Selanjutnya tahap cleaning ini juga termasuk tahap text processing. dimana tahap ini dilakukan untuk pemrosesan kata untuk dipisahkan satu per satu

```
[52] # Tokenize
tokenized = df['comment'].apply(lambda x:x.split())
tokenized

0      [biasa, itu, lagi, ada, kuasa, galak, ntar, tu...
1      [pdi, pasang, aniesahok, aja, pasti, menang]
2      [presiden, mulut, mencla, mencle, terulang, la...
3      [inilah, dinasti, politik, penguasa, yang, pal...
4      [terlihat, jelas, suzan, ternyata, para, senlo...
...
1181     [pak, kowi, tolong, pikirkan, juga, nasib, bob...
1182     [buruk, rupa, cerain, dibelah, sontoloyo]
1183     [bisa, jeda, nanti, ditambah, minimal, 17, tahun...
1184     [klan, dinasti, pragmatis, everlasting, cepat,...
1185     [ditunggu2]
Name: comment, Length: 1186, dtype: object
```

Gambar 4.6 Kodingan untuk Tokenizing dan hasil

4.5 Pelabelan Sentimen

Selanjutnya tahap pelabelan sentimen. dimana tahap ini dilakukan untuk mengidentifikasi kata untuk dipisahkan satu per satu antara sentimen negatif, positif dan netral



username	comment	sentiment
@sarmilih1233	biasa itu lagi ada kuasa galak ntar turun jadi...	negatif
@tamamesta	pdi pasang anies-ahok aja pas5 menang	negatif
@christinesuparmi2365	presiden mulut mencela mencela terulang lagi men...	negatif
@christinesuparmi2365	silah dinasti politik penguasa yang paling ra...	negatif
@muhamadkafa8397	terlihat jelas sudah ternyata para senior seni...	negatif
@stevenahmad1474	Pak Kowi tolong pikirkan juga nasib Bobby Jan ...	negatif
@ali_hchanne6073	Buruk rupa cermin dibelah sontoloyo	negatif
@ali_hchanne6073	Bisa jadi nanti diubah minimal 17 tahun utk me...	negatif
@adorasigma	Klan dinasti pragmatis everlasting sat set sea...	negatif
@fitriarurulh4967	diunggu2	netral

Gambar 4.7 Hasil Pelabelan Sentiment

4.6 Perhitungan Naive Bayes

Selanjutnya, pemodelan dilakukan menggunakan Algoritma Naive Bayes untuk menghitung akurasi, presisi, dan recall.

	Precision	Recall	f1-score	support
negatif	0.56	0.95	0.70	116
netral	0.85	0.31	0.46	108
positif	0.00	0.00	0.00	13
accuracy			0.61	237
macro avg	0.47	0.42	0.39	237
weighted avg	0.66	0.61	0.55	237

Tabel 4.1 Hasil perhitungan Naïve Bayes

Hasil penelitian menunjukkan bahwa model Naive Bayes dapat dengan akurat mendeteksi sentimen negatif. 0,56 dan recall sebesar 0,95, namun kurang efektif dalam mendeteksi emosi netral dan positif dengan skor f1 sebesar 0,46 atau 0,00. Akurasi keseluruhan model adalah 0,61

5. KESIMPULAN DAN SARAN

Tujuan penelitian ini adalah untuk melihat bagaimana pengguna YouTube bertindak balas terhadap video yang membahas keputusan Mahkamah Agung yang diduga memungkinkan Kaesang Pangarep, putra bungsu Presiden Joko Widodo, mencalonkan diri sebagai Wakil Gubernur Jakarta. Penelitian ini mengkategorikan komentar-komentar menjadi sentimen positif, negatif, dan netral dengan menggunakan metode Naive Bayes, sebuah algoritma yang umum digunakan untuk klasifikasi teks berdasarkan probabilitas. Hasil penelitian menunjukkan bahwa model Naive Bayes cukup efektif dalam menemukan sentimen negatif dengan presisi sebesar 0,56 dan recall sebesar 0,95, tetapi kurang efektif dalam menemukan sentimen netral dan positif dengan skor F1 sebesar 0,95. Model ini memiliki

akurasi total 0,61. Sentimen masyarakat terhadap masalah ini menunjukkan bagaimana masyarakat merespons keputusan ini.

Disarankan untuk menggabungkan metode Naive Bayes dengan algoritma lain atau menggunakan metode pembelajaran kelompok untuk meningkatkan akurasi dalam mendeteksi sentimen netral dan positif. Ini dapat membantu mengimbangi kekurangan algoritma tertentu dengan kelebihan algoritma lainnya. Menggunakan dataset yang lebih besar dan beragam juga dapat membantu model menemukan pola sentimen yang lebih kompleks. Dengan mengumpulkan komentar dari berbagai video dan platform media sosial, hal ini dapat dicapai. Sebelum diolah oleh model, proses preprocessing data seperti stemming, lemmatization, dan penghapusan stopwords harus dilakukan dengan lebih baik. Analisis lebih lanjut tentang komponen yang mempengaruhi sentimen publik, seperti latar belakang sosial, demografi, dan afiliasi politik, dapat memberikan wawasan yang lebih mendalam tentang alasan di balik sentimen yang muncul. Selain sentimen positif, negatif, dan netral, model juga dapat dievaluasi untuk mengidentifikasi sentimen lain seperti campuran atau sentimen yang lebih khusus, seperti skeptis atau harapan. Analisis sentimen diharapkan dapat memberikan hasil yang lebih akurat dan bermanfaat tentang pemahaman publik tentang masalah politik tertentu dengan mempertimbangkan rekomendasi ini.

REFERENSI

- Agusta, L. (2009). Perbandingan Algoritma Stemming Porter dengan Algoritma Nazief & Adriani untuk Stemming Dokumen Teks Bahasa Indonesia. Konferensi Nasional Sistem dan Informatika.
- Dabas, P. K., & Nurdin, G. C. (2021). Analisis Komentar di Video Youtube menggunakan Hadoop. *Jurnal Nasional Komputasi dan Teknologi Informasi*.
- Harpizon, R. K. I., Rasyid, E. B. F. S. H. A., & Rasyid, S. (2022). Analisis Sentimen Komentar di YouTube Tentang Ceramah Ustadz Abdul Somad Menggunakan Algoritma Naïve Bayes. *Jurnal Nasional Komputasi dan Teknologi Informasi*.
- Muthia, D. A. (2017). Analisis Sentimen pada Review Restoran dengan Teks Bahasa Indonesia Menggunakan Naive Bayes. *Jurnal Ilmu Pengetahuan dan Teknologi Komputer*.
- Rahman, W. D. A. (2017). Online News Classification Using Multinomial Naive Bayes. *Jurnal Teknologi dan Informasi*.
- Septian, T. F. A. N. F. I. K. N. J. A. (2019). Analisis Sentimen Pengguna Twitter Terhadap Polemik Persepakbolaan Indonesia Menggunakan Pembobotan TF-IDF dan K-Nearest Neighbor. *Journal of Intelligent Systems and Computation*.

- Styawati, N. H. A. R. I. A. Y. R. (2021). Analisis Sentimen Masyarakat Terhadap Program Kartu Prakerja pada Twitter Dengan Metode Support Vector Machine. Jurnal TRANSFORMATIKA.
- Surojudin, A. S. d. N. (2020). Analisis dan Perbandingan Stemming Algoritma Porter dengan Algoritma Ahmad Yusoff Sembok dalam Dokumen Teks Bahasa Indonesia. Seminar Nasional Teknologi Informasi dan Komunikasi STI&K.
- Syahril Dwi Prasetyo, F. N., & Shofa Shofiah Hilabi, S. (2023). Analisis Sentimen Relokasi Ibukota Nusantara Menggunakan Algoritma Naïve Bayes dan KNN. Jurnal KomtekInfo.
- Ulil Albab, Y. K. P. M. N. F. M. (2023). Optimization of the Stemming Technique on Text Preprocessing President. Jurnal TRANSFORMATIKA.
- Zhafira, B. R., & Irwanda, I. D. F. (2021). Analisis Sentimen Kebijakan Kampus Merdeka Menggunakan Naive Bayes dan Pembobotan TF-IDF Berdasarkan Komentar pada Youtube. Jurnal Teknologi Informasi dan Ilmu Komputer (JUST-SI).