

# Analisis Sentimen Pengobatan Supranatural Di Media Sosial X Menggunakan Metode Naïve Bayes

Feliex Zebaoth <sup>1,\*</sup>, Wahyu Teja Kusuma <sup>2</sup>, M. Syauqi Haris <sup>3</sup>

<sup>1,2,3</sup>Program studi Informatika, Institut Teknologi, Sains, dan Kesehatan RS.DR.Soepraoen Kesdam V/BRW, Malang, Indonesia

Email: <sup>1,\*</sup>felixzebaoth@gmail.com, <sup>2</sup>wtkusuma@itsk-soepraoen.ac.id, <sup>3</sup>haris@itsk-soepraoen.ac.id

<sup>\*)</sup> Email Penulis Utama

**Abstrak**– Praktik pengobatan supranatural di Indonesia, yang sering dianggap sebagai alternatif atau pelengkap pengobatan medis modern, memicu perdebatan di tengah kemajuan ilmu kedokteran, dengan persepsi masyarakat yang sangat bervariasi. Penelitian ini diperlukan untuk memahami persepsi tersebut, karena memengaruhi literasi kesehatan dan kebijakan dalam konteks budaya yang beragam. Menggunakan teknik crawling pada pengumpulan data, 1794 tweet dikumpulkan dari platform X (2016–2025), diikuti dengan pembersihan data untuk menghapus gangguan, duplikat, dan unggahan tidak relevan serta pre-processing data. Analisis sentimen dilakukan dengan model dari algoritma Naïve Bayes. Model mencapai akurasi 82%, dengan precision, recall, dan F1-score masing-masing 86%, 82%, dan 84% untuk sentimen Positif, serta 77%, 83%, dan 80% untuk sentimen Negatif. *Confusion matrix* menunjukkan 112 *True Positives*, 85 *True Negatives*, 18 *False Positives*, dan 25 *False Negatives*. Analisis word cloud mengungkapkan sentimen Positif mencerminkan opini yang kontra terhadap pengobatan supranatural (“shaman”, “ruqyah”, “doctor”, “medical”), sedangkan sentimen Negatif menunjukkan dukungan terhadap pengobatan supranatural (“believe”, “healer”, “spiritual”). Studi ini diharapkan dapat menyuguhkan perspektif yang lebih jelas tentang sikap masyarakat terhadap praktik supranatural, Selain mengetahui persepsi masyarakat terdapat pengobatan supranatural penelitian ini juga dapat dimanfaatkan untuk usulan atau kritik konstruktif dengan menerapkan strategi komunikasi yang lebih efektif dan pendekatan empatik bagi tenaga medis serta pembuat kebijakan guna meningkatkan literasi kesehatan. Dengan memahami faktor yang memengaruhi kepercayaan pada pengobatan supranatural, temuan ini diharapkan bisa mendukung perancangan regulasi kesehatan yang inklusif dan berbasis data, sehingga layanan kesehatan ilmiah dapat lebih diterima oleh masyarakat di segala tingkatan.

**Kata Kunci:** Pengobatan Supranatural, Analisis Sentimen, Naive bayes, Media Sosial X, Klasifikasi Teks.

**Abstract**– Supernatural treatment practices in Indonesia, often regarded as alternatives or complements to modern medical treatment, spark debate amid advancements in medical science, with public perceptions varying widely. This study is essential to understand these perceptions, as they influence health literacy and policy within a diverse cultural context. Using data crawling techniques, 1,794 tweets were collected from the X platform (2016–2025), followed by data cleaning to remove noise, duplicates, and irrelevant posts, as well as data pre-processing. Sentiment analysis was conducted using a model based on the Naïve Bayes algorithm. The model achieved an accuracy of 82%, with precision, recall, and F1-score of 86%, 82%, and 84% for Positive sentiments, and 77%, 83%, and 80% for Negative sentiments, respectively. The confusion matrix indicated 112 True Positives, 85 True Negatives, 18 False Positives, and 25 False Negatives. Word cloud analysis revealed that Positive sentiments reflect opposition to supernatural treatment (“shaman,” “ruqyah,” “doctor,” “medical”), while Negative sentiments indicate support for supernatural treatment (“believe,” “healer,” “spiritual”). This study aims to provide a clearer perspective on public attitudes toward supernatural practices. Beyond understanding public perceptions of supernatural treatment, this research can be utilized to propose or offer constructive criticism by implementing more effective communication strategies and empathetic approaches for healthcare professionals and policymakers to enhance health literacy. By understanding the factors influencing belief in supernatural treatment, these findings are expected to support the design of inclusive, data-driven health regulations, ensuring that scientific healthcare services are more widely accepted across all societal levels.

**Keywords:** Supernatural Healing, Sentiment Analysis, Naive Bayes, Social Media X, Text Classification.

## 1. PENDAHULUAN

Di era pesatnya kemajuan teknologi, sering ditemukan beberapa fenomena dan fakta-fakta yang menunjukkan adanya kepercayaan masyarakat Indonesia terhadap supranatural [1]. Menurut Kamus Besar Bahasa Indonesia (KBBI), supranatural adalah hal yang bersifat mistis, ajaib, gaib, atau adikodrati yang melampaui batas alamiah

dalam konteks ini praktik supranatural sering dikaitkan dengan peran dukun sebagai pelaku pengobatan tradisional yang memanfaatkan kekuatan gaib.

Dalam empat tahun terakhir, fenomena supranatural terlihat di berbagai provinsi, seperti praktisi penyembuhan supranatural di Jawa Timur [2] [3], guru spiritual maut di Bogor Jawa Barat [4], ritual untuk mengusir *covid-19* di Sragen Jawa Tengah [5], pengobatan yang mengatas namakan budaya Dayak di Kupang NTT [6], dan paranormal sekaligus entertainer di dunia hiburan pada DKI Jakarta [7]. Kepercayaan kuat masyarakat terhadap hal mistis menyebabkan mistisifikasi berkembang masif, dengan banyaknya konten bermuatan mistis yang mendapat perhatian luas di media sosial [2]. Sebagian paranormal bahkan mengaitkan praktik mereka dengan agama, menegaskan bahwa kesembuhan berasal dari Tuhan, berbeda dengan dokter yang mengandalkan teknologi medis canggih. Namun, meskipun teknologi medis semakin maju, tantangan baru seperti munculnya penyakit yang belum ditemukan penawarnya tetap ada [8]. Selain itu kepercayaan yang berlebihan terhadap pengobatan supranatural dibandingkan metode medis menimbulkan kekhawatiran di masyarakat, karena berisiko menghambat diagnosis dan perawatan penyakit yang seharusnya ditangani dengan pendekatan medis yang telah terbukti secara ilmiah [9].

Dalam perkembangan informasi di era digital, media sosial telah berfungsi sebagai salah satu fasilitas utama dalam menyebarkan informasi dan menyampaikan opini. Salah satu yang paling populer adalah *Twitter*, yang berganti nama menjadi *X* pada 22 Juli 2023 [10]. Di Indonesia, platform ini memiliki sekitar 18,45 juta pengguna aktif pada tahun 2022, digunakan untuk berbagi informasi, berdiskusi, dan menyampaikan pandangan terkait berbagai topik. Hal ini menciptakan volume data yang besar di platform tersebut, menjadikannya sumber yang relevan untuk menganalisis opini masyarakat berdasarkan kata kunci yang spesifik [11]. Proses menganalisis opini atau komentar dalam bentuk teks termasuk dalam cabang ilmu analisis sentimen [10]. Analisis sentimen merupakan pendekatan komputasi yang mempelajari opini dan emosi yang terkandung dalam teks [12]. Beberapa penelitian sebelumnya, seperti analisis sentimen terhadap isu kesehatan di media sosial [13], analisis sentimen program jaminan kesehatan nasional menggunakan *multiclass svm* [14]. Namun, belum ada penelitian yang secara khusus menganalisis persepsi masyarakat terhadap supranatural melalui pendekatan serupa.

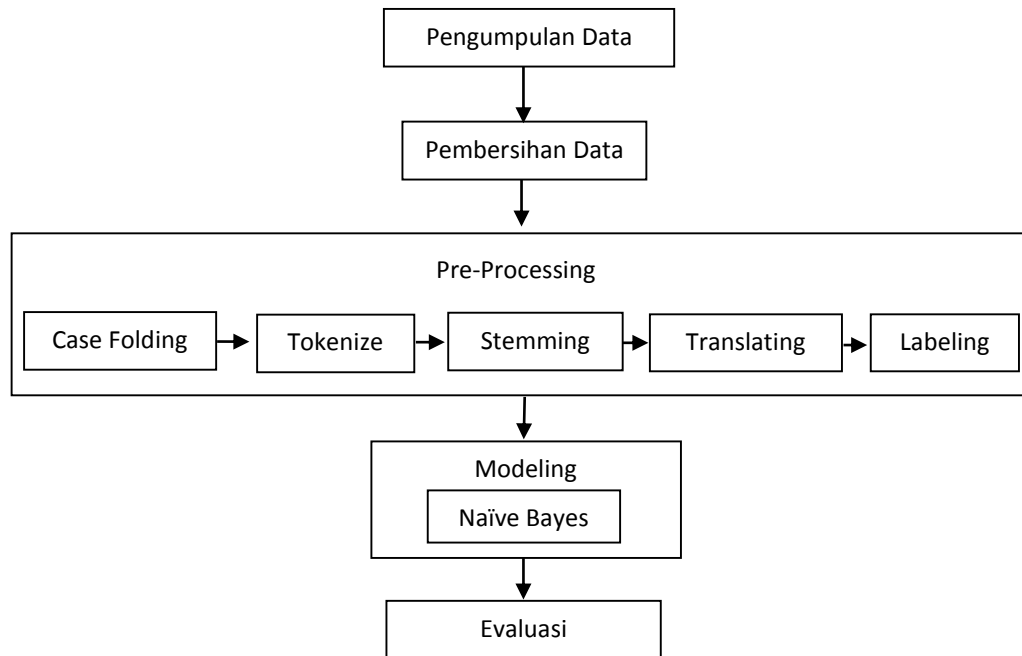
Untuk menggali lebih dalam pola opini masyarakat tersebut, pembelajaran mesin dapat digunakan apabila tersedia data sebagai input, dengan menganalisis kumpulan data dalam jumlah besar [15]. Pola-pola tersebut dapat ditemukan, dalam konteks analisis sentimen, algoritma *Naïve Bayes* merupakan salah satu teknik yang sering dimanfaatkan dalam analisis sentimen untuk mengelompokkan teks ke dalam kelas positif, negatif, atau netral [16]. Metode ini menghitung probabilitas berdasarkan kata-kata dalam dokumen untuk menentukan sentimennya [17] [18]. Oleh karena itu, penelitian ini menggunakan metode *Naïve Bayes* untuk menganalisis sentimen masyarakat terkait pengobatan supranatural di *Twitter*, dengan data dari kurun waktu 12 November 2016 hingga 13 Januari 2025. Semakin baik data dimodelkan semakin akurat dan andal pula hasil yang dihasilkan [19]. Pemilihan alat dan teknik yang tepat untuk memodelkan data bergantung pada jenis data, tugas yang dicapai, dan keterbatasan komputasi [19]. Sehingga penggunaan *Naïve Bayes* dalam analisis ini diharapkan mampu menawarkan klasifikasi sentimen yang efektif dan relevan.

Penelitian ini bertujuan untuk mengeksplorasi opini masyarakat mengenai pengobatan supranatural melalui analisis sentimen di media sosial. Dengan menerapkan metode *Naïve Bayes*, penelitian ini mengelompokkan sentimen ke dalam kategori positif dan negatif untuk memahami pola persepsi publik. Studi ini diharapkan dapat menyuguhkan perspektif yang lebih jelas tentang sikap masyarakat terhadap praktik supranatural serta menekankan pentingnya pendekatan empatik dalam layanan kesehatan. Selain itu, riset ini turut berfokus pada identifikasi faktor sosial dan ekonomi yang memengaruhi kecenderungan masyarakat dalam memilih pengobatan non-medis. Selain mengetahui persepsi masyarakat terhadap pengobatan supranatural penelitian ini juga dapat dimanfaatkan untuk mengembangkan strategi komunikasi yang lebih efektif dan pendekatan empatik bagi tenaga medis serta pembuat kebijakan guna meningkatkan literasi kesehatan. Dengan memahami faktor sosial-ekonomi yang memengaruhi kepercayaan pada pengobatan supranatural, temuan ini mendukung perancangan regulasi kesehatan yang inklusif dan berbasis data, sehingga layanan kesehatan ilmiah dapat lebih diterima oleh masyarakat.

## 2. METODE PENELITIAN

Metode penelitian penilaian sentimen terhadap pengobatan supranatural di media sosial *X* di ilustrasikan oleh gambar 1 Data dikumpulkan melalui *crawling* dengan kata kunci “pengobatan supranatural,” “percaya dukun,”

“pengobatan mistik,” dan “pengobatan gaib”. diikuti *cleansing* untuk menghapus noise seperti tanda baca dan emoji. Prapemrosesan meliputi *case folding*, *tokenization*, *stopword removal*, *stemming*, *translating*, labeling, Evaluasi. Naïve Bayes diterapkan dengan aturan teorema Bayes yang mengasumsikan independensi kata untuk menghitung probabilitas sentimen. Model dilatih dengan data berlabel untuk mengklasifikasi unggahan, lalu dievaluasi menggunakan akurasi dan *precision*. Hasil klasifikasi digunakan untuk menyimpulkan persepsi masyarakat.



Gambar 1. Tahap Penelitian

## 2.1 Pengumpulan Data

Langkah awal dalam proses ini adalah pengumpulan data menggunakan teknik *crawling*. Pada tahap ini, data mentah diambil dari platform Twitter, berupa unggahan dan tweet yang mengandung kata kunci seperti “pengobatan supranatural,” “percaya dukun,” “pengobatan mistik,” dan “pengobatan gaib.” Data mentah yang diperoleh akan menjadi sumber utama untuk dilakukan analisis sentimen [20].

## 2.2 Pembersihan Data

Pada tahap selanjutnya merupakan proses pembersihan data dengan teknik *cleansing* data tahap awal untuk menghilangkan gangguan pada data. Langkah-langkahnya meliputi penghapusan tanda baca, simbol, menghapus *tweet* duplikat atau *spam*, *drop* kolom meta data dan memperbaiki ejaan kata [21].

## 2.3 Pre-processing Data

Pengolahan awal atau *preprocessing* merupakan tahapan penting dalam pengolahan data sebelum dianalisis lebih lanjut. Tahap ini bertujuan untuk menyaring data dari elemen-elemen yang tidak diperlukan atau tidak relevan, mengganggu, seperti duplikasi, tanda baca, atau kata-kata kosong. Dengan melakukan *preprocessing*, data yang dihasilkan menjadi lebih terstruktur, konsisten, dan siap digunakan dalam proses pemodelan atau analisis berikutnya [22].

### a) Case folding

*Case Folding* adalah tahapan yang mengubah seluruh teks menjadi huruf kecil untuk menjamin konsistensi data sebelum analisis sentimen, Dengan standarisasi ini, data teks menjadi lebih seragam, memudahkan tahapan berikutnya [20]

b) **Tokenize**

*Tokenize* adalah merupakan proses membagi teks menjadi bagian-bagian kecil, seperti kata atau frasa yang dikenal sebagai token, sebagai langkah awal dalam pemrosesan.ekstraksi fitur sentimen setiap tweet. yang menjadi dasar analisis [23].

c) **Stemming**

*Stemming* adalah menyederhanakan kata-kata dalam token ke bentuk dasarnya. untuk mengurangi variasi leksikal dalam analisis sentimen. Teknik ini, yang umum dalam pengolahan bahasa alami, memastikan kata dengan akar yang sama dianggap identik [20].

d) **Translating**

*Translating* adalah melakukan perubahan pada teks yang awalnya berbahasa indonesia akan diubah ke bahasa inggris. Penerjemahan teks ke dalam bahasa yang seragam memastikan bahwa semua data yang ditujukan untuk pelabelan mematuhi format yang konsisten. Praktik ini mengurangi potensi kesalahan yang mungkin timbul dari perbedaan bahasa [24].

e) **Labeling**

*Labeling* adalah merupakan tahapan memberikan label sentimen positif atau negatif pada teks tweet yang telah diolah melalui tahap sebelumnya, untuk mempersiapkan data pelatihan bagi model serta mendukung klasifikasi sentimen yang konsisten [20].

2.4 **Modeling**

Modeling adalah merupakan tahap di mana dilakukan pengembangan model atau algoritma untuk analisis data. Pada proses ini, algoritma dirancang dan dilatih agar mampu mengenali pola dalam data teks. Tahapan ini menghasilkan sebuah model yang mampu mengelompokkan teks ke dalam kategori sentimen tertentu, seperti sentimen positif maupun negatif [25].

3. **HASIL DAN PEMBAHASAN**

3.1 **Pengumpulan data**

Studi ini mengumpulkan data berasal dari jejaring sosial X menggunakan teknik crawling dengan alat Tweet Harves, yang menghimpun unggahan berdasarkan kata kunci seperti “pengobatan supranatural,” “percaya dukun,” “pengobatan mistik,” dan “pengobatan gaib”. Proses pengumpulan data berlangsung dari 12 November 2016 hingga 13 Januari 2025, menghasilkan 1.794 tweet yang disimpan dalam bentuk file CSV guna dilakukan analisis lanjutan. Tahapan ini, sebagaimana diilustrasikan pada Gambar 2, merupakan langkah awal dalam metode penelitian untuk menganalisis sentimen masyarakat terhadap pengobatan supranatural.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1		conversat	created_at	favorit	full_text	id_str	imag	in_reply_to_scre	lang	location	quote_co	reply_retw	tweet_url		user_id_str	username
2	0	1.88E+18	Sun Jan 12 15:12	0	gue males lanjut jc	1.88E+18			in	Indonesi	0	0	0	https://x.com/nadizyy/	1.01161E+18	nadizyy
3	1	1.88E+18	Sun Jan 12 12:00	0	@nightride Ga se	1.88E+18		nightride	in	Indonesi	0	0	0	https://x.com/Minamo	1.14526E+18	Minamo_Ken
4	2	1.88E+18	Sat Jan 11 23:49	2	@tambakansor @e	1.88E+18		tambakansor	in		0	0	0	https://x.com/Agus639	1.85578E+18	Agus6390Arnold
5	3	1.88E+18	Sat Jan 11 14:08	0	@Yunaqshin @1tii	1.88E+18		Yunaqshin	in	ðŸ’ã€	0	0	0	https://x.com/minjdo	3391928299	minjdo
6	4	1.88E+18	Sat Jan 11 13:28	0	@PH1B0YYY gaperi	1.88E+18		PH1B0YYY	in		0	1	1	https://x.com/sorasudi	1.66037E+18	sorasudiro

Gambar 2. hasil Pengumpulan data

3.2 **Cleansing Data**

Tahap *cleansing* dilakukan setelah pengumpulan 1.794 *tweet* untuk menghilangkan noise seperti mention (@username), tautan, emotikon, tanda baca, dan karakter khusus dari kolom *full\_text*, serta menghapus kolom

metadata seperti *conversation\_id\_str*, *favorite\_count*, *id\_str*, *image\_url*, *in\_reply\_to\_screen\_name*, *lang*, *location*, *quote\_count*, *reply\_count*, *retweet\_count*, *tweet\_url*, dan *username*, sehingga hanya *full\_text* yang disisakan. Proses ini, yang merupakan langkah penting dalam pengolahan bahasa alami, memastikan bahwa data teks bersih dan terfokus pada konten yang relevan untuk analisis sentimen menggunakan Naïve Bayes. Perbandingan data sebelum dan sesudah *cleansing*, yang menunjukkan teks yang lebih terstruktur, disajikan pada Tabel 1.

**Tabel 1** Sebelum Dan Sesudah Cleansing Data

Sebelum Cleansing	Sesudah Cleansing
@nightride Ga semua RS kali yaudah lu pada berobat ke dukun aja kalo gitu. Syukur deh kalo mau pada bayar umum sih	Ga semua RS kali yaudah lu pada berobat ke dukun aja kalo gitu. Syukur deh kalo mau pada bayar umum sih
@tambakansor @eexoo_7 Orang kek lu mending berobat ke dukun aja gih gausah ke RS dimantrain diludahin cuih paling sembuh.	Orang kek lu mending berobat ke dukun aja gih gausah ke RS dimantrain diludahin cuih paling sembuh.

### 3.3 Pre-Processing Data

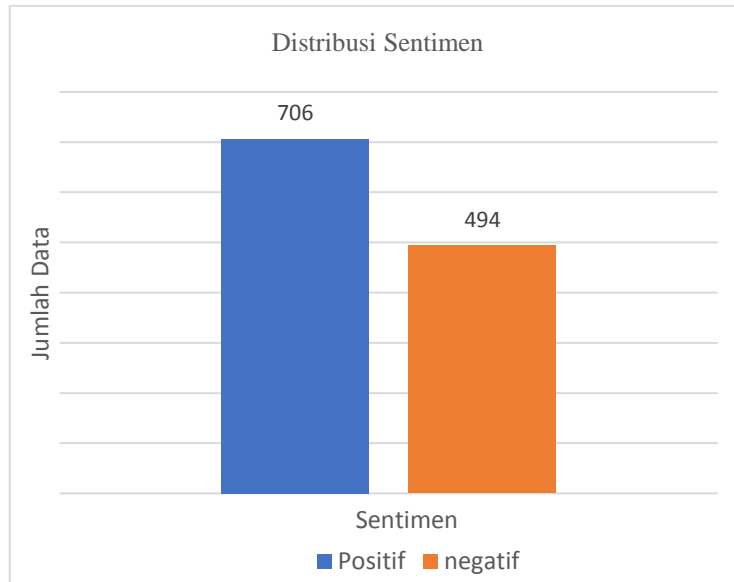
Setelah *cleansing* untuk menghilangkan noise dan metadata dari *tweet* tentang pengobatan supranatural, tahap pra-pemrosesan dilanjutkan dengan serangkaian langkah untuk mempersiapkan teks dalam kolom *full\_text* untuk analisis sentimen, meliputi *case folding* untuk menyeragamkan huruf menjadi huruf kecil, *tokenization* untuk menguraikan teks menjadi kata-kata, *stemming* untuk mengembalikan kata ke bentuk dasar (misalnya, "berlari" menjadi "lari"), *translating* untuk menerjemahkan teks berbahasa Indonesia ke bahasa Inggris menggunakan alat penerjemah otomatis, dan *labelling* untuk menentukan sentimen; langkah penerjemahan diperlukan karena *TextBlob*, *library* Python untuk pemrosesan teks alami, dioptimalkan untuk analisis sentimen dalam bahasa Inggris berdasarkan pola kata dan frasa. Teks terjemahan (*tweet\_english*) dianalisis menggunakan *TextBlob* untuk menghitung polaritas (-1 hingga 1) dan subjektivitas (0 hingga 1), dengan polaritas  $\geq 0$  diklasifikasikan sebagai "Positif" (misalnya, "luar biasa") dan  $< 0$  sebagai "Negatif" (misalnya, "kecewa"), di mana polaritas = 0 dianggap Positif untuk menyederhanakan klasifikasi biner, meskipun mencerminkan ekspresi tanpa sentimen kuat. Subjektivitas mengukur tingkat opini, dengan nilai mendekati 0 untuk fakta dan 1 untuk opini. Hasil analisis disusun dalam Tabel 2, mencakup sebelum dan sesudah pada tahapan *pre-processing* data, sentimen, polaritas, dan subjektivitas, dengan dua data acak dipilih untuk mewakili sentimen positif dan negatif guna menganalisis pola opini masyarakat terhadap pengobatan supranatural.

**Tabel 2** Sebelum Dan Sesudah Pre-Processing Data

Sebelum <i>preprocessing</i>	Setelah <i>pre-processing</i>	Sentimen	Polarity	Subjectivity
bawa ke dukun aja sudah kaga mempan pengobatan konvensional tu	Just take him to a shaman, conventional treatment doesn't work	Negatif	-0.142	0.357
berobat pake bpjs gratis sementara kalo ke dukun dan herbal2an doang bisa habis puluhan juta itupun jarang yang sampe sembuh meskipun proses pengobatan kanker di rs panjang tapi banyak yang berhasil sembuh	seek treatment using bpjs for free while if you go to shamans and herbs doang can run out of tens of millions even then it is rare to recover even though the cancer treatment process in rs is long but many have managed to recover	Positif	0.287	0.65

Setelah *cleansing* menghasilkan *tweet* bebas noise dan *pre-processing* menstandarkan *full\_text* melalui *case folding*, *tokenization*, *stemming*, dan *translating* ke bahasa Inggris, *TextBlob* digunakan untuk melabeli *tweet* berdasarkan polaritas terhadap 1200 data. Distribusi sentimen awal mengalami ketidakseimbangan signifikan,

dengan 706 *tweet* berlabel Positif (78,5%) dan 193 *tweet* berlabel Negatif (21,5%). Ketidakseimbangan ini menyebabkan bias untuk inputan model nantinya terhadap kelas mayoritas (Positif). Untuk mengatasi ketidakseimbangan, teknik *random oversampling* diterapkan pada kelas minoritas menggunakan pustaka *imblearn* dengan rasio 70% dari kelas mayoritas. Proses ini meningkatkan jumlah *tweet* Negatif dari 193 menjadi 494, sementara jumlah *tweet* Positif tetap 706, sebagaimana dirangkum dalam Tabel 2 Diagram batang pada Gambar 3 memvisualisasikan distribusi sentimen ini, mengungkap dominasi sentimen positif terhadap pengobatan supranatural, yang memfasilitasi pemahaman awal tentang opini masyarakat.



Gambar 3 Total distribusi sentimen.

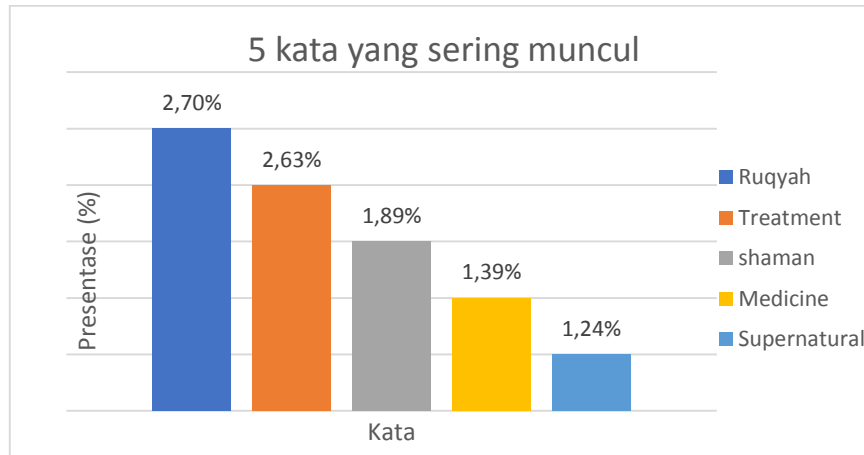
Gambar diagram batang di atas menunjukkan analisis sentimen yang dilakukan setelah tahap pre-processing menggunakan tools *TextBlob* dan dilakukan teknik *random oversampling* menunjukkan bahwa mayoritas opini publik mengenai pengobatan supranatural memiliki sentimen positif, dengan total 706 *tweet* berlabel positif dan 494 *tweet* berlabel negatif. Temuan ini menunjukkan bahwa sebagian besar pengguna media sosial memiliki pandangan yang kontra terhadap praktik pengobatan supranatural. Di sisi lain, jumlah opini negatif yang lebih kecil mengindikasikan bahwa masih adanya kepercayaan terhadap metode pengobatan ini setelah visualisasi jumlah sentimen tahap selanjutnya ialah Visualisasi word cloud dilakukan untuk mengidentifikasi kata kunci dominan dalam setiap kelas sentimen.



Gambar 4 Hasil dari tahap pre-processing divisualisasikan melalui Word Cloud.

Visualisasi Word Cloud pada Gambar 4 ilustrasi di atas mencerminkan persebaran istilah yang sering ditemukan pada percakapan terkait pengobatan supranatural, baik dengan sentimen positif maupun negatif. Pada word cloud sentimen positif (kiri), kata-kata seperti "treatment", "ruqyah", "shaman", dan "supernatural" muncul dengan ukuran besar, menandakan frekuensi tinggi penggunaannya dalam konteks yang bernada mendukung. Hal ini menunjukkan bahwa sebagian masyarakat masih memandang pengobatan alternatif seperti ruqyah dan praktik supranatural sebagai bentuk penanganan kesehatan yang layak diperhatikan. Sebaliknya, pada word cloud sentimen negatif (kanan), kata "shaman", "doctor", dan "people" menjadi yang paling dominan, menandakan

adanya kritik, keraguan, atau pandangan negatif terhadap pengobatan supranatural maupun konvensional. Kata-kata seperti "crazy", "sick", dan "afraid" juga turut muncul, mencerminkan nada skeptis dan kekhawatiran terhadap praktik-praktik tersebut. Visualisasi ini memberikan gambaran bahwa pengobatan supranatural masih menjadi topik perdebatan, dengan opini yang terbagi antara yang mendukung dan yang meragukan. Untuk memperkuat pemahaman terhadap persepsi publik ini, analisis frekuensi kata lebih lanjut dapat dilakukan untuk mengidentifikasi kata-kata dengan kemunculan tertinggi dalam dataset, gambar di bawah menyajikan lima kata dengan frekuensi tertinggi dalam bentuk diagram batang.



Gambar 5 Presentase kata yang sering muncul.

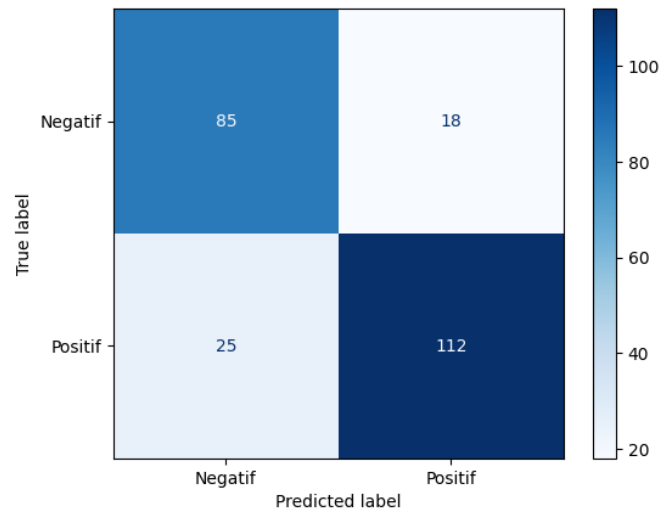
Diagram batang pada gambar 5 di atas menampilkan lima istilah dengan jumlah kemunculan tertinggi dalam dataset yang telah melewati tahap *pre-processing*. Kata "ruqyah" menjadi yang paling dominan dengan muncul sebanyak 345 kali (2,70%), diikuti oleh "treatment" (pengobatan) yang tercatat sebanyak 336 kali (2,63%). Sementara itu, istilah "shaman" (dukun) ditemukan sebanyak 242 kali (1,89%), "medicine" (obat) muncul dalam 178 unggahan (1,39%), dan "supernatural" (supranatural) tercatat sebanyak 159 kali (1,24%). Distribusi kata ini menunjukkan bahwa perbincangan mengenai pengobatan supranatural banyak berkisar pada istilah-istilah tersebut, mengindikasikan adanya hubungan antara metode penyembuhan tradisional dan pendekatan medis modern dalam diskusi publik. Fase berikutnya ialah menerapkan model pengklasifikasian sentimen dengan memakai algoritma Naïve Bayes, yang dikenal efektif dalam menganalisis teks. Model ini langsung dipanggil dan dilatih menggunakan dataset yang telah diproses untuk mengukur performa klasifikasi sentimen. Serta dilakukan tahap evaluasi pada model, ini bertujuan untuk mengetahui seberapa tinggi performa model dapat mengenali dan mengklasifikasikan sentimen berdasarkan data yang telah diberi label sebelumnya menggunakan *TextBlob*. langkah berikutnya mengubah yaitu teks *tweet* menjadi vektor frekuensi kata menggunakan *CountVectorizer* untuk menghasilkan representasi numerik bagi pelatihan model, Data yang telah disampling ulang dibagi menjadi data latih (80%, 960 data) dan data uji (20%, 240 data). Model Naïve Bayes dilatih pada data latih, dan Evaluasi performa model berdasarkan classification report, yang tertera pada Tabel 3.

Tabel 3 Evaluasi Model Naïve Bayes

	classification report			
	Precision	Recall	F1-score	support
Negative	77	83	80	103
Positive	86	82	84	137
Accuracy			82	240
Macro avg	82	82	82	240
Weighted avg	82	82	82	240

Tabel 3 menunjukkan bahwa model mencapai tingkat akurasi sebesar 82%, yang mencerminkan performa klasifikasi yang cukup baik secara keseluruhan. Model mampu mengenali sentimen positif dengan *precision* sebesar 86%, *recall* 82%, dan *F1-score* 84%, yang menunjukkan bahwa model cukup andal dalam mendeteksi opini positif dari data. Sementara itu, untuk sentimen negatif, model memperoleh *precision* sebesar 77%, *recall* 83%, dan *F1-score* 80%. Hasil ini menunjukkan bahwa meskipun performa terhadap sentimen negatif sedikit

lebih rendah dibandingkan sentimen positif, perbedaan kinerjanya tidak terlalu mencolok, dan model tetap mampu menangani kedua kelas secara seimbang. Untuk melihat lebih detail bagaimana model mengklasifikasikan masing-masing sentimen, dilakukan analisis menggunakan *confusion matrix*. Matriks ini memperlihatkan sejauh mana model mampu mengidentifikasi sentimen dengan benar serta menunjukkan jenis kesalahan yang terjadi selama proses klasifikasi. Untuk memahami lebih lanjut bagaimana model mengklasifikasikan sentimen positif dan negatif, dilakukan analisis menggunakan *confusion matrix*. Tabel berikut menyajikan distribusi prediksi model terhadap kelas sentimen, yang mencerminkan sejauh mana model berhasil mengidentifikasi sentimen dengan benar serta kesalahan yang masih terjadi dalam proses klasifikasi.



Gambar 6. *Confusion Matrix*

Confusion matrix, sebagaimana ditunjukkan pada Gambar 6, digunakan untuk mengevaluasi efektivitas model klasifikasi sentimen terhadap dataset pengobatan supranatural setelah penerapan Random Oversampling dengan rasio 70%. Matriks ini mengukur jumlah prediksi benar dan salah untuk dua kelas sentimen, yaitu Positif dan Negatif, pada 240 data uji, memberikan gambaran performa model dalam mengenali opini publik di media sosial.

Distribusi prediksi model berdasarkan confusion matrix adalah sebagai berikut:

1. True Positives (TP): Model berhasil mengklasifikasikan 112 data berlabel Positif dengan benar sebagai Positif, menunjukkan kemampuan yang kuat dalam mengenali opini yang mendukung pengobatan supranatural.
2. True Negatives (TN): Sebanyak 85 data berlabel Negatif diklasifikasikan dengan tepat sebagai Negatif, mencerminkan keberhasilan model dalam mendeteksi opini skeptis terhadap pengobatan supranatural.
3. False Positives (FP): Terdapat 18 data berlabel Negatif yang salah diklasifikasikan sebagai Positif (Type I error), mengindikasikan sejumlah kecil kesalahan dalam mengidentifikasi opini negatif.
4. False Negatives (FN): Sebanyak 25 data berlabel Positif keliru diprediksi sebagai Negatif (Type II error), menunjukkan adanya tantangan dalam mendeteksi semua opini positif.

Interpretasi confusion matrix menunjukkan bahwa model memiliki performa yang cukup baik, dengan jumlah prediksi benar (TP: 112, TN: 85) yang signifikan dibandingkan kesalahan (FP: 18, FN: 25). Penerapan Random Oversampling, yang meningkatkan jumlah data Negatif dari 193 menjadi 494 untuk mengatasi ketidakseimbangan awal (706 Positif vs. 193 Negatif), berkontribusi pada peningkatan recall kelas Negatif (0,83, sebagaimana ditunjukkan dalam laporan klasifikasi). Hal ini menunjukkan bahwa model lebih mampu mendeteksi opini negatif dibandingkan tanpa resampling, mengurangi bias terhadap kelas mayoritas (Positif). Meskipun demikian, adanya 25 False Negatives dan 18 False Positives mengindikasikan ruang untuk perbaikan, terutama dalam mengurangi kesalahan klasifikasi pada kedua kelas. Secara keseluruhan, *confusion matrix* ini menggambarkan kemampuan model yang seimbang dalam mengklasifikasikan sentimen Positif dan Negatif setelah penyeimbangan data. wawasan penting tentang dinamika opini publik terhadap pengobatan supranatural di media sosial.

## 4. KESIMPULAN

Penelitian ini mengkaji opini masyarakat Indonesia terhadap pengobatan supranatural melalui analisis sentimen pada 1794 tweet menggunakan algoritma Naïve Bayes, mencapai akurasi keseluruhan 82% dengan 706 sentimen positif dan 494 sentimen negatif. Analisis word cloud mengungkapkan bahwa sentimen Positif, dengan kata kunci seperti “*shaman*”, “*ruqyah*”, “*doctor*”, dan “*medical*”, mencerminkan opini yang kontra dengan pengobatan supranatural. Sebaliknya, sentimen Negatif, dengan kata seperti “*believe*”, “*healer*”, dan “*spiritual*”, menunjukkan dukungan terbatas berbasis keyakinan budaya. Meskipun model Naïve Bayes efektif mendeteksi sentimen Positif, keterbatasan dalam mengidentifikasi sentimen Negatif, sebagian akibat kesalahan klasifikasi (25 False Negatives, 18 False Positives), menunjukkan tantangan dalam menangkap hubungan kontekstual antar kata pada data minoritas. Temuan ini menegaskan perlunya pendekatan empatik dalam layanan kesehatan dan strategi komunikasi inklusif oleh tenaga medis untuk meningkatkan literasi kesehatan dan mengatasi skeptisisme masyarakat. Untuk meningkatkan akurasi, disarankan menggunakan Support Vector Machine (SVM), yang unggul dalam menangani fitur tekstual berdimensi tinggi dan hubungan non-linear, atau Recurrent Neural Network (RNN), yang mampu menangkap konteks berurutan dalam teks, sehingga lebih efektif untuk tweet dengan pola bahasa kompleks. Selain itu, memperluas dataset, memasukkan sentimen netral, menerapkan ekstraksi topik, dan mengonversi emoticon dapat memberikan wawasan lebih komprehensif tentang persepsi masyarakat terhadap pengobatan supranatural, mendukung regulasi kesehatan berbasis data yang mempertimbangkan faktor sosial-ekonomi.

## UCAPAN TERIMAKASIH

Penulis menyampaikan apresiasi yang sebesar-besarnya kepada lembaga pendidikan, institusi penelitian, serta pihak-pihak terkait yang telah memberikan dukungan, bimbingan, dan fasilitas dalam pelaksanaan studi ini. Harapan besar penulis adalah Temuan dalam penelitian ini diharapkan bermanfaat bagi pembaca serta menjadi kontribusi yang berarti dalam bidang keilmuan yang dikaji.

## REFERENCES

- [1] A. R. Togatorop, A. V. Sinaga, and J. A. Tan, “Mysticism and Traditional Medicine: A Study of Christian Theology on Mysticism and Traditional Medicine and its Reflections for Christians Today,” *J. Relig. Socio-Cultural*, vol. 4, no. 2, pp. 171–198, 2023, doi: <https://doi.org/10.46494/psc.v18i2.216>.
- [2] M. F. Ni’am, “Dari Mistis Ke Bisnis: Praktik Okultisme Gus Syamsudin di Youtube,” *Al-Adyan J. Relig. Stud.*, vol. 3, no. 2, pp. 142–155, 2022, [Online]. Available: <https://ejournal.uinib.ac.id/jurnal/index.php/aladyan/article/view/4997>
- [3] R. : F. P. A. E. : F. Pradita, “Sama-sama Dukun Seperti Gus Samsudin, Ningsih Tinampi Juga Suka ‘Sembuhkan’ Orang, Kesaktiannya Didapat dari Sosok Dedengkot Dukun,” *tvonenews.com*, 2022, [Online]. Available: <https://www.tvonenews.com/lifestyle/trend/58267-sama-sama-dukun-seperti-gus-samsudin-ningsih-tinampi-juga-suka-semuhkan-orang-kesaktiannya-didapat-dari-sosok-dedengkot-dukun?page=all>
- [4] Rizky Adha Mahendra, “Guru Spiritual Maut di Bogor Ngaku Bisa Sembuhkan Segala Penyakit,” *news.detik.com*. [Online]. Available: <https://news.detik.com/berita/d-6828645/guru-spiritual-maut-di-bogor-ngaku-bisa-semuhkan-segala-penyakit>
- [5] Levi Larassaty, “Heboh! Dukun Gelar Ritual di Alun-alun Sragen untuk Mengusir Covid-19,” *health.grid.id*, 2020, [Online]. Available: <https://health.grid.id/read/352083904/heboh-dukun-gelar-ritual-di-alun-alun-sragen-untuk-mengusir-covid-19?page=all>
- [6] Antonius un and Taulin, “Pengobatan Ida Dayak di Kupang,” *gatra.com*, 2023, [Online]. Available: <https://www.gatra.com/news-571557-hukum-pengobatan-ida-dayak-di-kupang-hoaks-korban-transfer-uang-muka.html>
- [7] Rintan puspita sari, “Sosok Ki Joko Bodo, Lulusan S1 yang Memilih Menekuni Profesi Paranormal,” *kompas.com*, 2022, [Online]. Available: <https://www.kompas.com/hype/read/2022/11/22/123302166/sosok-ki-joko-bodo-lulusan-s1-yang-memilih-menekuni-profesi-paranormal?page=all>
- [8] T. Firdaus and N. B. Abdul, “Kepercayaan Masyarakat Terhadap Pengobatan Supranatural Di Desa Batang Kabupaten Bulukumba,” *J. Socius Educ. ( JSE )*, pp. 201–209, 2023, doi: 10.0505/jse.v.
- [9] N. Jannah and Y. Y. Prasetyawan, “Available Online at : <http://ejournal.undip.ac.id/index.php/anuva> Pengobatan Alternatif di Desa Suwawal : Analisis Persepsi dan Perilaku Informasi Masyarakat Abstrak,” vol. 8, no. 3, pp. 409–424, 2024, doi: 10.14710/anuva.8.3.409-424.

- [10] F. Adi Artanto, "Analisis Sentimen Opini Publik terhadap Fenomena Bunuh Diri Mahasiswa di Twitter Menggunakan Algoritma Naive Bayes," vol. 4, no. 1, pp. 70–77, 2024, doi: 10.54259/satesi.v4i1.2908.
- [11] R. Hidayat, M. Fikry, F. Yanto, and E. Pandu Cynthia, "Penerapan Naïve Bayes Classifier dalam Klasifikasi Sentimen Publik di Twitter terhadap Puan Maharani," *JUKI J. Komput. dan Inform.*, vol. 6, no. 1.
- [12] M. N. Muttaqin and I. Kharisudin, "Analisis Sentimen Pada Ulasan Aplikasi Gojek Menggunakan Metode Support Vector Machine dan K Nearest Neighbor," *UNNES J. Math.*, vol. 10, no. 2, pp. 22–27, 2021, [Online]. Available: <http://journal.unnes.ac.id/sju/index.php/ujm>
- [13] F. A. Hermawati, N. S. Fatonah, and H. A. Mangambali, "Aplikasi Analisis Sentimen Isu Kesehatan di Media Sosial dengan Metode Convolutional Neural Network Berbasis Web," *J. Eksplora Inform.*, vol. 12, no. 2, pp. 120–128, 2024, doi: 10.30864/eksplora.v12i2.1012.
- [14] N. Gusti, A. Dasriani, L. Ahmad, G. Pariandi, and I. M. Y. Dharma, "Analisis Sentimen Program Jaminan Kesehatan Nasional Menggunakan Multiclass Support Vector Machine," vol. 6, no. 1, pp. 20–30, 2025, doi: doi.org/10.37148/bios.v6i1.136.
- [15] M. S. Haris, A. N. Khudori, and W. T. Kusuma, "Perbandingan Metode Supervised Machine Learning untuk Prediksi Prevalensi Stunting di Provisi Jawa Timur," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 9, no. 7, p. 1571, 2022, doi: 10.25126/jtiik.2022976744.
- [16] F. Septianingrum and A. S. Y. Irawan, "Metode Seleksi Fitur Untuk Klasifikasi Sentimen Menggunakan Algoritma Naive Bayes: Sebuah Literature Review," *J. Media Inform. Budidarma*, vol. 5, no. 3, p. 799, 2021, doi: 10.30865/mib.v5i3.2983.
- [17] T. Hartati, R. T. Sohadi, E. Tohidi, and E. Wahyudin, "Penerapan Algoritma Naive Bayes pada Analisis Sentimen Ulasan Aplikasi Whoosh – Kereta Cepat Di Google Play Store," vol. 6, no. 1, pp. 244–249, 2024.
- [18] D. S. Nugroho, I. F. Hanif, M. A. Hasbi, F. Fredianto, A. M. Saputra, and R. Zildjian, "Analisis Sentimen Dugaan Pelanggaran Pemilu 2024 Berdasarkan Tweet Menggunakan Algoritma Naïve Bayes Classifier," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 4, no. 3, pp. 1169–1176, Jul. 2024, doi: 10.57152/malcom.v4i3.1496.
- [19] W. T. K. Widyastuti Andriyani, Mochammad Anshori, Dwi Normawati, Risqy Siwi Pradini, Mohamad Zaenudin, Muhammad Iqbal Harisuddin, M. Syauqi Haris, Astuty, Anna Angela Sitingjak, *Matematika Pada Kecerdasan Buatan*. CV.Tohar Media, 2024. [Online]. Available: <https://books.google.co.id/books?id=cEYhEQAAQBAJ&lpg=PP1&ots=q4tcQ4YGTc&dq=info%3AbBM2Qj5sgHkJ%3Ascholar.google.com&lr&pg=PP1#v=onepage&q&f=false>
- [20] M. F. Saleh and R. Imanda, "Public Sentiment Analysis of the Free Meal Program : A Comparison of Naive Bayes and Support Vector Machine Methods on the Twitter ( X ) Social Media Platform," vol. 9, no. 1, pp. 131–139, 2025, doi: 10.30871/jaic.v9i1.8895.
- [21] D. Musfiroh, U. Khaira, P. E. P. Utomo, and T. Suratno, "Analisis Sentimen terhadap Perkuliahan Daring di Indonesia dari Twitter Dataset Menggunakan InSet Lexicon," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 1, no. 1, pp. 24–33, 2021, doi: 10.57152/malcom.v1i1.20.
- [22] I. S. Arfan, S. Fauziah, and I. Nawangsih, "Sentiment Analyst of Cyber Bullying in X Using Naïve Bayes Algorithm Analisa Sentimen Terhadap Cyber Bullying di X Menggunakan Algoritma Naïve Bayes," vol. 4, no. October, pp. 1411–1419, 2024.
- [23] Aditya Quantano Surbakti, Regiolina Hayami, and Januar Al Amien, "Analisa Tanggapan Terhadap Psbb Di Indonesia Dengan Algoritma Decision Tree Pada Twitter," *J. CoSciTech (Computer Sci. Inf. Technol.)*, vol. 2, no. 2, pp. 91–97, 2021, doi: 10.37859/coscitech.v2i2.2851.
- [24] E. Mailoa, "Perbandingan Beberapa Algoritma Machine Learning Dalam Analisis Sentimen Terkait Pemilihan Presiden RI 2024," 2024, doi: 10.35889/jutisi.v13i2.1980.
- [25] M. Daffa, A. Fahreza, A. Luthfiarta, M. Rafid, M. Indrawan, and A. Nugraha, "JOURNAL OF APPLIED COMPUTER SCIENCE AND TECHNOLOGY ( JACOST ) Analisis Sentimen : Pengaruh Jam Kerja Terhadap Kesehatan Mental Generasi Z," vol. 5, no. 1, pp. 16–25, 2024, doi: /10.52158/jacost.v5i1.715.