



# UNIVERSITAS BHAYANGKARA JAKARTA RAYA

## FAKULTAS TEKNIK

Kampus I: Jl. Harsono RM No.67, Ragunan, Pasar Minggu, Jakarta Selatan 12550

Telepon: (021) 27808121 - 27808882

Kampus II: Jl. Raya Perjuangan, Marga Mulya, Bekasi Utara, Jawa Barat

Telepon: (021) 88955882 Fax.: (021) 88955871

Web: [www.ubharajaya.ac.id/ft/](http://www.ubharajaya.ac.id/ft/). Email: [ft@ubharajaya.ac.id](mailto:ft@ubharajaya.ac.id)

### SURAT TUGAS

Nomor: ST/608/X/2022/FT-UBJ

Pertimbangan : Bahwa dalam rangka Penelitian dan Pengabdian kepada Masyarakat Universitas Bhayangkara Jakarta Raya sebagai penunjang Tri Dharma Perguruan Tinggi di Universitas Bhayangkara Jakarta, maka untuk itu perlu dikeluarkan Surat Tugas.

Dasar : 1. Keputusan Mendiknas RI Nomor : 184/V/2001, tanggal 23 Nopember 2001, tentang Pedoman, Pengendalian dan Pembinaan Program Diploma, Sarjana dan Pascasarjana di Perguruan Tinggi.  
2. Keputusan Dirjen Dikti Nomor : 034/Dikti/Kep/2002, tanggal 3 Juli 2002, tentang Perubahan dan Peraturan Tambahan Pedoman Pengawasan Pengendalian dan Pembinaan Program Diploma, Sarjana dan Pascasarjana di Perguruan Tinggi.  
3. Rencana Kerja dan Anggaran Pembelanjaan Universitas Bhayangkara Jakarta Raya Tahun 2022/2023.

### DITUGASKAN

Kepada : Dr. Tyastuti Sri Lestari, S.Si., M.M.  
Dosen Fakultas Teknik

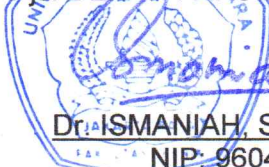
NIDN : 0327036701

Untuk : 1. Bertugas untuk melaksanakan Penelitian dengan judul "Analisis Sentimen Terhadap Kementerian Perdagangan pada Sosial Media Twitter Menggunakan Metode Naive Bayes".  
2. Kegiatan dilakukan pada Oktober – Desember 2022.  
3. Melaporkan hasil pelaksanaan kegiatan tersebut secara tertulis kepada Dekan Fakultas Teknik Universitas Bhayangkara Jakarta Raya.  
4. Melaksanakan perintah ini dengan penuh rasa tanggung jawab.

Dikeluarkan di : Bekasi

Pada Tanggal : 3 Oktober 2022

Pjs. DEKAN FAKULTAS TEKNIK



Dr. ISMANIAH, S.Si., M.M.

NIP. 9604028

Paraf:

1. Ka. Prodi TID...

# Analisis Sentimen Terhadap Kementerian Perdagangan Pada Sosial Media Twitter Menggunakan Metode Naïve Bayes

*Sentiment Analysis Of The Ministry Of Trade On Twitter Social Media Using Naïve Bayes Method*

Fefbiansyah Hasibuan<sup>1</sup>, Wowon Priatna<sup>2</sup>, Tyastuti Sri Lestari<sup>3</sup>

<sup>1,2</sup>informatika, universitas bhayangkara jakarta raya

<sup>3</sup>Teknik Industri, universitas bhayangkara jakarta raya

E-mail: <sup>1</sup>fefbiansyah.hasibuan18@mhs.ubharajaya.ac.id, <sup>2</sup>wowon.priatna@dsn.ubhrajaya.ac.id,

<sup>3</sup>tyas@ubharajaya.ac.id

## Abstrak

Penelitian ini untuk mendapatkan opini masyarakat di media sosial *Twitter* terkait tentang Kementerian Perdagangan Republik Indonesia dengan mengimplementasikan algoritma *Naïve Bayes* untuk melakukan analisis sentimen terhadap opini yang ada. Penelitian ini bertujuan untuk menganalisis opini masyarakat di media sosial terhadap Kementerian Perdagangan Republik Indonesia mengenai kelangkaan minyak goreng dengan menggunakan *Naïve Bayes*. Data yang digunakan pada penelitian ini adalah postingan *tweet* yang diberikan masyarakat yang ditujukan dan berkaitan kepada Kementerian Perdagangan Republik Indonesia yang diambil sebanyak 1000 *tweet*. Metode pengumpulan data dilakukan dengan cara crawling menggunakan *access token api key* yang di dapat dari *twitter developer*. Setelah data didapat maka dilakukan *text processing* agar mempermudah dalam proses analisis. Hasil analisis pada penelitian ini menggunakan algoritma *naïve bayes* adalah dengan nilai akurasi sebesar 89,24%. Perbandingan persentase didapatkan 84,02% tanggapan yang diberikan masyarakat bernilai positif dan 15,98% bernilai negatif.

Kata kunci: Analisis Sentimen, Naïve Bayes, Kementerian Perdagangan Republik Indonesia, Tweet, Klasifikasi

## Abstract

*The topic raised in this research is discussing the issue of public opinion on Twitter related to the Ministry of Trade of the Republic of Indonesia by implementing the Naïve Bayes algorithm to conduct sentiment analysis on existing opinions. This study aims to analyze public opinion on social media against the Ministry of Trade of the Republic of Indonesia using the Naïve Bayes algorithm. The data used in this study are tweets given by the public aimed at and related to the Ministry of Trade of the Republic of Indonesia which were taken as many as 1000 tweets. The data collection method is carried out by crawling using the access token api key which is obtained from the twitter developer. After the data is obtained, text processing is carried out to facilitate the analysis process. The results of the analysis in this study using the Naive Bayes algorithm is with an accuracy value of 89.24%. Comparison of percentages obtained 84.02% of the responses given by the community are positive and 15.98% are negative.*

Keywords: Sentiment Analysis, Naïve Bayes, Ministry of Trade of the Republic of Indonesia, Tweet, Clasification.

## 1. PENDAHULUAN

Kementerian Perdagangan RI adalah suatu lembaga pemerintah yang sudah ada sejak tanggal 2 September 1945 dan pada awalnya dikenalkan dengan nama Kementerian Kemakmuran. Yang memiliki wewenang menangani pekerjaan dalam bidang perdagangan dalam menolong presiden untuk mengendalikan lembaga pemerintahan negara bagian perdagangan [1]. Badan bagian Humas Kementerian Perdagangan Republik Indonesia selalu memberikan informasi berkaitan mengenai kegiatan Kementerian Perdagangan melalui beberapa akun media sosialnya seperti Twitter, Facebook, Youtube, dan Instagram. Akun media sosial Twitter milik Kementrian Perdagangan Republik Indonesia sering membagikan informasi mengenai perkembangan perdagangan tentang produk dalam negeri, kebijakan pemerintah, kesuksesan pemerintah, membagikan informasi mengenai perdagangan, investasi, dan industri. Informasi dan kebijakan yang diberikan oleh Kementrian Perdagangan sering mendapatkan kritikan dan pujian. Salah satu informasi dari Kementrian Perdagangan yang membuat banyak komentar mengenai kelangkaan minyak goreng dipasaran ketika masih dijual berpatokan pada harga tertentu berbanding terbalik ketika harga subsidi dicabut ketersediannya melimpah dipasaran. Dikarenakan banyak keputusan yang bertentangan dengan keinginan rakyat dan dinilai gagal dalam menjaga kesetabilan harga bahan pokok. komentator banyak ditujukan masyarakat kepada akun twitter kementrian perdagangan terkait kebijakan kelangkaan minyak goreng bersubsidi dipasaran membuat banyak masyarakat yang kesusahaan mencari barang tersebut, Tetapi ketika dicabutnya subsidi terhadap minyak goreng maka ketersediaan barang tersebut tersedia. Kebijakan dari kementrian tersebut menimbulkan ragam komentar yang beropini negatif atau positif sehingga masuk kedalam daftar topik yang populer dalam media *Social Twitter*. Untuk mengetahui mengklasifikasikan komentar termasuk opini positif dan negatif perlu dilakukan analisis sentiment agar mendapatkan informasi yang optimal untuk memanfaatkan media sosial untuk kepentingan publik [2].

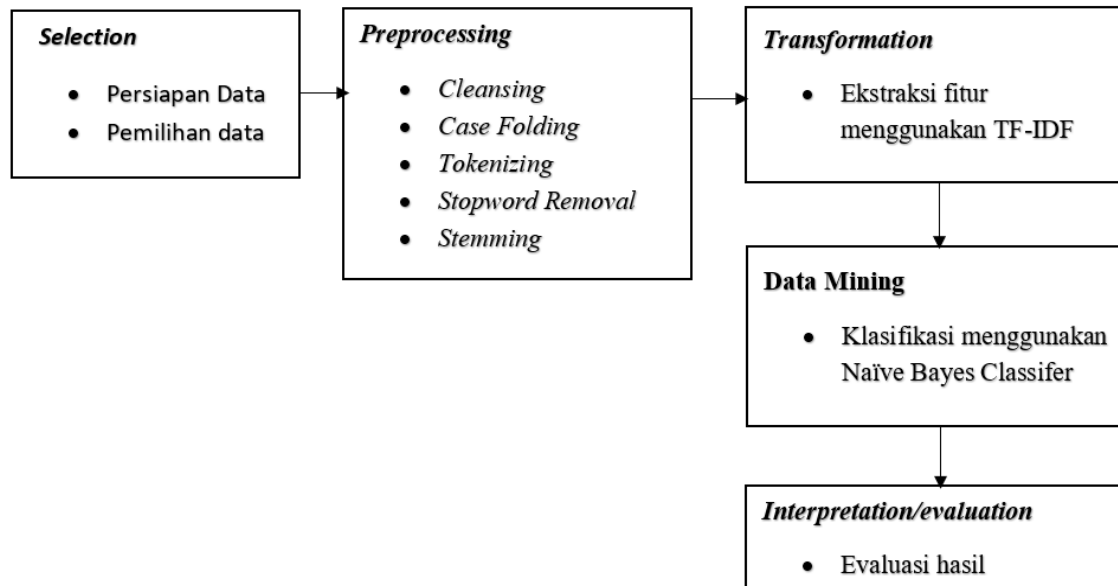
Analisis sentimen atau *opinion mining* adalah Sebagian implementasi dari *text mining* yang mengangkat opini yang disampaikan oleh masyarakat, sentimen, penilaian, sikap, dan emosi pada entitas dan atribut yang berfokus adalah mengekstraksi, mengidentifikasi dan meneliti dan menemukan karakteristik sentimen atau perilaku seseorang yang disampaikan di media sosial terhadap masalah yang timbul [3] [4] [5]. Beberapa penelitian menggunakan algoritma *Naïve Bayes* dalam melakukan analisis sentimen diantaranya penelitian[6] melakukan analisis sentimen untuk mendapatkan klasifikasi informasi pengguna produk BPJS, sentimen analisis untuk mengetahui persepsi pelanggan dalam menggunakan berbagai sumber online [7], analisis sentimen dari pendapat masyarakat di media *Twitter* berupa tanda emosional yang digunakan untuk mengukur persepsi pelanggan secara komputasi [8] dan analisis sentimen proses komputasi mengidentifikasi dan mengkategorikan pendapat yang diungkapkan dalam sepotong teks, terutama untuk menentukan apakah sikap penulis terhadap topik tertentu, produk dan lainnya adalah positif, negatif, atau netral [9].

Beberapa penelitian terkait sentimen analisis adalah sentimen analisis Masyarakat terhadap Layanan *Shopee food* menggunakan algoritma *Naïve Bayase* menghasilkan opini positif dengan nilai akurasi 90,62% [10], Sentimen untuk kartu prakerja dengan algoritma *Naïve Bayase* menghasilkan opini positif menunjukkan manfaat untuk pekerja dengan akurasi 80,1% [11], sentimen untuk mengetahui opini masyarakat terhadap pemilihan gubernur DKI Jakarta menggunakan *Naïve Bayase* menghasilkan opini positif dengan nilai akurasi sebesar 96% [5], analisis sentimen tentang pidato presiden Jokowi dengan *preprocessing* normalisasi dan *stemming* menggunakan metode *Naïve Bayes* dan *SVM* akurasi 89,2655% [12], analisis sentimen mengenai vaksinasi menggunakan *Naïve Bayes* dan *Social Network Analysis (SNA)* akurasi sebesar 92% [13], Melakukan sentimen dengan mereview produk kosmetik dengan algoritma *naïve bayes* memiliki akurasi 90,50% [14], analisa sentimen Twitter menggunakan klasifikasi *Naïve Bayes* pada promosi wisata Museum Sangiran Kabupaten Sragen dengan nilai akurasi 87,91% [15].

Berdasarkan permasalahan dan penelitian terdahulu belum pernah ada analisis sentimen untuk mengetahui opini masyarakat terhadap kelangkaan minyak goreng yang bersumber dari media sosial Kementerian Perdagangan Republik Indonesia, maka tujuan penelitian ini adalah untuk menganalisis opini masyarakat Indonesia mengenai kelangkaan minyak goreng dengan menggunakan metode Naïve Bayes sebagai algoritma untuk klasifikasi data.

## 2. METODE PENELITIAN

Desain penelitian untuk melakukan sentimen analisis dapat dijabarkan pada gambar 1.



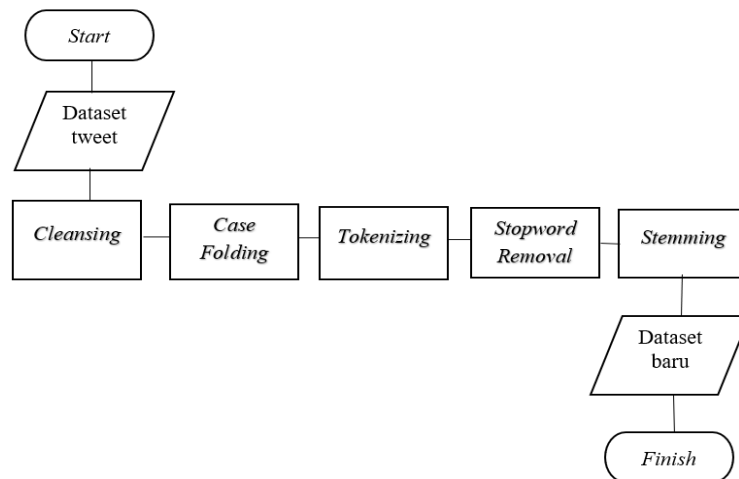
Gambar 1 Desain Penelitian

### 2.1 Selection

Fase ini merupakan fase persiapan dan seleksi data. Data yang telah diambil dari *Twitter* akan dilakukan proses pemilihan untuk melanjutkan ke tahap *preprocessing*[16]. Data yang diambil merupakan opini yang diberikan terhadap Kementerian Perdagangan Republik Indonesia. Data diambil dari tanggal, 3 Januari 2022 sampai 31 maret 2022.

### 2.2 Preprocessing

Tahap ini adalah tahap yang terpenting untuk tahap selanjutnya, yaitu untuk mereduksi atribut-atribut yang kurang berdampak pada proses klasifikasi. Berikut ini adalah beberapa langkah-langkah *preprocessing* yang disajikan pada gambar 2.



Gambar 2 Preprocessing

Tahap *preprocessing* berguna untuk membersihkan kata agar mempermudah dalam proses analisis. Tahapan-tahapan pada proses *preprocessing* antara lain sebagai berikut:

- a. *Case Folding* berfungsi untuk merubah kalimat yang didalamnya terdapat huruf besar menjadi huruf kecil semua dengan menggunakan bantuan *sintax* dari python yaitu *str.lower(.)*
- b. *Cleansing* bermanfaat untuk membersihkan data tweet yang berisi angka, tanda baca, link, hastag, mention seperti [!'"#\$%&'\_~] dengan menggunakan manfaat *library* dari python yaitu *library string*.
- c. Tahap *tokenizing*  
Berfungsi untuk memisahkan kata-kata yang digabungkan dengan spasi ke dalam setiap kata yang dikumpulkan dalam sebuah array sehingga sebuah kalimat terpecah menjadi kata-kata. Pada tahap ini, kita akan menggunakan *library ppython, nltk.tokenize*.
- d. *Stopword removal* bermanfaat untuk menghilangkan kata-kata yang tidak mempengaruhi pada proses analisis sentimen.
- e. *Stemming* pada tahapan ini berfungsi untuk merubah kata yang berimbuhan menjadi kata dasar untuk mempermudah pembobotan.

a. *Naïve Bayes*

*Naïve Bayes* adalah metode klasifikasi yang tepat untuk mencari nilai probabilitas tertinggi secara cepat, sederhana, dan akurat. Berikut merupakan perhitungan dari *naïve bayes*[17].

$$V_{map} = \operatorname{argmax} V_j \{ \text{Positif, Negatif} \} p(v_j) \pi P(x_i | V_j) \quad (1)$$

Keterangan :

- $V_j$  : Kategori tweet  $j = 1, 2, \dots$ dst.  
Misal  $j_1 =$  kelas positif,  $j_2$  kelas negatif
- $P(x_i | v_j)$  : probabilitas  $x_i$  pada kategori  $V_j$
- $P(V_j)$  : Probalitas dari  $V_j$

Untuk  $P(V_j)$  dan  $P(x_i | V_j)$  dihitung pada saat pelatihan dimana persamaannya adalah sebagai berikut:

$$P(V_j) = \frac{|docs\ j|}{|contoh|} \quad (2.2)$$

$$P(x_i | V_j) = \frac{nk + 1}{n + |kosakata|} \quad (2.3)$$

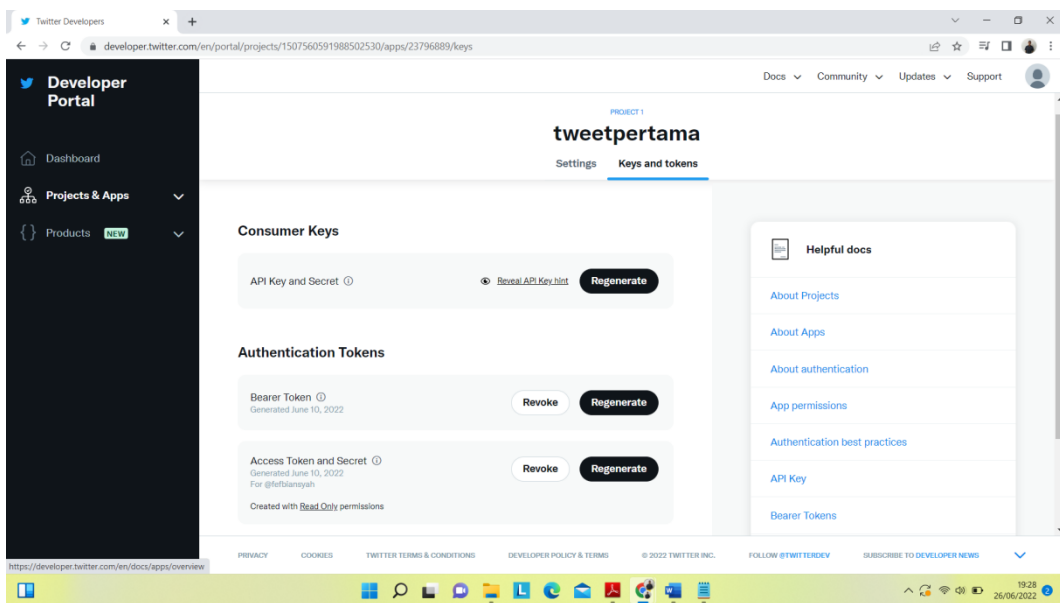
Keterangan :

- | docs j | : jumlah dokumen setiap kategori
- | contoh | : jumlah dokumen dari semua kategori
- nk : jumlah kemunculan frekuensi dari setiap kata
- n : jumlah frekuensi kemunculan kata setiap kategori
- | kosakata | : jumlah semua kata dari semua kategori

### 3. HASIL DAN PEMBAHASAN

#### 3.1 Implementasi

Implementasi pada langkah ini dimulai dengan tahap pengumpulan data. Dengan cara melakukan perhitungan emosional manual bobot kata untuk mendapatkan klasifikasi secara otomatis. Tahapan ini dilakukan pengambilan data melalui tahap *crawling* pada *twitter* yang sebelumnya telah mendapatkan *api key* dari twitter developer.



Gambar 3 Access Token API

#### 3.2 Tahap *Preprocessing*

Tahap *preprocessing* berguna untuk membersihkan kata agar mempermudah dalam proses analisis. Tahap *preprocessing* terdiri dari beberapa tahap antara lain:

Tabel 1 Tahap *Preprocessing*

Tahapan	Hasil
Data Tweet asli	terima Kasih banyak atas dukuNgan dan Upaya @KEMENDAG yang mendukUng sektOr https://perdagangan di tanah air.
<i>Cleansing</i>	terima Kasih banyak atas dukuNgan dan Upaya KEMENDAG yang mendukUng sektOr perdagangan di tanah air.
<i>Case Folding</i>	terima kasih banyak atas dukungan dan upaya kemendag yang mendukung sektor perdagangan di tanah air.
<i>Tokenizing</i>	'terima' 'kasih' 'banyak' 'atas' 'dukungan' 'dan' 'upaya' 'kemendag' 'yang' 'mendukung' 'sektor' 'perdagangan' 'tanah' 'air'.
<i>Stopword Removal</i>	terimakasih banyak dukungan dan upaya kemendag mendukung sektor perdagangan tanah air
<i>Stemming</i>	terimakasih banyak dukung dan upaya kemendag sektor dagang tanah air

3.3 Data mining

Pada tahap ini, *Naïve Bayes* digunakan untuk melakukan klasifikasi. Adapun implementasi klasifikasi dari data uji dapat dilihat pada tabel 2.

Tabel 2 Data Training

Tweet	Fitur	Kategori
Tweet1	Terimakasih kepada kemendag	P
Tweet2	Bangga dengan produk lokal	P
Tweet3	Dukung kemendag promosikan produk lokal ke pasar dunia	P
Tweet4	Kecewa kepada kemendag gagal stabilkan bahan pokok	N
Tweet5	Dirjen dukung mafia bahan pokok	N
Tweet6	Pejabat membuat masyarakat kecewa dan sengsara	N

Dari data pada tabel 2, model probabilitas dibuat seperti pada persamaan berikut:  
Kata “terimakasih”:

$$P(a_{\text{terimakasih}} | V_{\text{positif}}) = \frac{1 + 1}{14 + 23} = \frac{2}{37}$$

$$P(a_{\text{terimakasih}} | V_{\text{negatif}}) = \frac{0 + 1}{18 + 23} = \frac{1}{41}$$

Kata “kepada”:

$$P(a_{\text{kepada}} | V_{\text{positif}}) = \frac{1 + 1}{14 + 23} = \frac{2}{37}$$

$$P(a_{\text{kepada}} | V_{\text{negatif}}) = \frac{1 + 1}{18 + 23} = \frac{2}{41}$$

Kata “kemendag”:

$$P(a_{\text{kemendag}} | V_{\text{positif}}) = \frac{2 + 1}{14 + 23} = \frac{3}{37}$$

$$P(a_{\text{kemendag}} | V_{\text{negatif}}) = \frac{1 + 1}{18 + 23} = \frac{2}{41}$$

Jika dikonversi ke tabel, maka probabilitas dari setiap kata dalam data uji pada analisis ini seperti akan terlihat pada tabel 3.

Tabel 3 Penghitungan Probabilitas Tweet1

Kategori	$P(a_i   V_j)$			
	$P(V_j)$	Terimakasih	Kepada	Kemendag
P	$\frac{1}{2}$	$\frac{2}{37}$	$\frac{2}{37}$	$\frac{2}{37}$
N	$\frac{1}{2}$	$\frac{2}{41}$	$\frac{2}{41}$	$\frac{1}{41}$

Tabel 4 Penghitungan Probabilitas Tweet2

Kategori	$P(a_i   V_j)$				
	$P(V_j)$	Bangga	Dengan	Produk	Lokal
P	$\frac{1}{2}$	$\frac{2}{37}$	$\frac{2}{37}$	$\frac{2}{37}$	$\frac{2}{37}$
N	$\frac{1}{2}$	$\frac{1}{41}$	$\frac{1}{41}$	$\frac{1}{41}$	$\frac{1}{41}$

Hasil perhitungan probabilitas digunakan sebagai model probabilitas selanjutnya dan sebagai data acuan untuk menentukan data uji proses Testing. Proses uji di sini, penulis menggunakan data uji seperti yang ditunjukkan pada tabel 5 untuk memberikan contoh cara menghitung data uji.

Tabel 5 Uji Data Testing

Tweet	Fitur	Kategori
Tweet 7	Terimakasih telah meningkatkan perdagangan	?
Tweet 8	Pemeliharaan mafia dilingkungan pemerintah kemendag	?
Tweet 9	Kementerian yang membuat kecewa masyarakat dan sengsara	?

$$\begin{aligned}
 P(Tweet7|V_{positif}) &= P(a_{terimakasih} | V_{positif}) \times \\
 &P(a_{meningkatkan} | V_{positif}) \times P(a_{perdagangan} | V_{positif}) \times P(positif) \\
 &= \frac{2}{37} \times 1 \times 1 \times 1 \times \frac{1}{2} = 0,027
 \end{aligned}$$

$$\begin{aligned}
 P(Tweet7|V_{negatif}) &= P(a_{terimakasih} | V_{negatif}) \times \\
 &P(a_{meningkatkan} | V_{negatif}) \times P(a_{perdagangan} | V_{negatif}) \times P(v_{negatif}) \\
 &= \frac{2}{41} \times 1 \times 1 \times \frac{1}{2} = 0,02439
 \end{aligned}$$

$$\begin{aligned}
 &P(Tweet8|V_{positif}) \\
 &= P(a_{pemeliharaan} | V_{positif}) \times P(a_{mafia} | V_{positif}) \times P(a_{dilingkungan} | V_{positif}) \times \\
 &P(a_{pemerintahan} | V_{positif}) \times P(a_{kemendag} | V_{positif}) \times P(positif) \\
 &= 1 \times \frac{1}{37} \times 1 \times 1 \times \frac{3}{37} \times \frac{1}{2} = 0,001095 \\
 &P(Tweet8|V_{negatif}) \\
 &= P(a_{pemeliharaan} | V_{negatif}) \times P(a_{mafia} | V_{negatif}) \times P(a_{dilingkungan} | V_{negatif}) \times \\
 &P(a_{pemerintahan} | V_{negatif}) \times \\
 &P(a_{kemendag} | V_{negatif}) \times P(negatif) \\
 &= 1 \times \frac{2}{41} \times 1 \times 1 \times \frac{2}{41} \times \frac{1}{2} = 0,00118976
 \end{aligned}$$

$$\begin{aligned}
 &P(Tweet9|V_{positif}) \\
 &= P(a_{kementrian} | V_{positif}) \times P(a_{membuat} | V_{positif}) \times \\
 &P(a_{kecewa} | V_{positif}) \times P(a_{masyarakat} | V_{positif}) \times P(a_{sengsara} | V_{positif}) \times \\
 &P(positif) \\
 &= 1 \times \frac{1}{37} \times \frac{1}{37} \times \frac{1}{37} \times \frac{1}{2} = 0,0000098145
 \end{aligned}$$

$$\begin{aligned}
 &P(Tweet9|V_{negatif}) \\
 &= P(a_{kementerian} | V_{negatif}) \times P(a_{membuat} | V_{negatif}) \times P(a_{kecewa} | V_{negatif}) \times \\
 &P(a_{masyarakat} | V_{negatif}) \times P(a_{sengsara} | V_{negatif}) \times P(positif) \\
 &= 1 \times \frac{2}{41} \times \frac{2}{41} \times \frac{2}{41} \times \frac{1}{2} = 0,000055296
 \end{aligned}$$

Setelah melakukan penghitungan probabilitas untuk setiap data uji, maka dapat didapatkan hasil seperti yang tercantum pada tabel 6.

Tabel 6 Hasil Probabilitas Data Testing

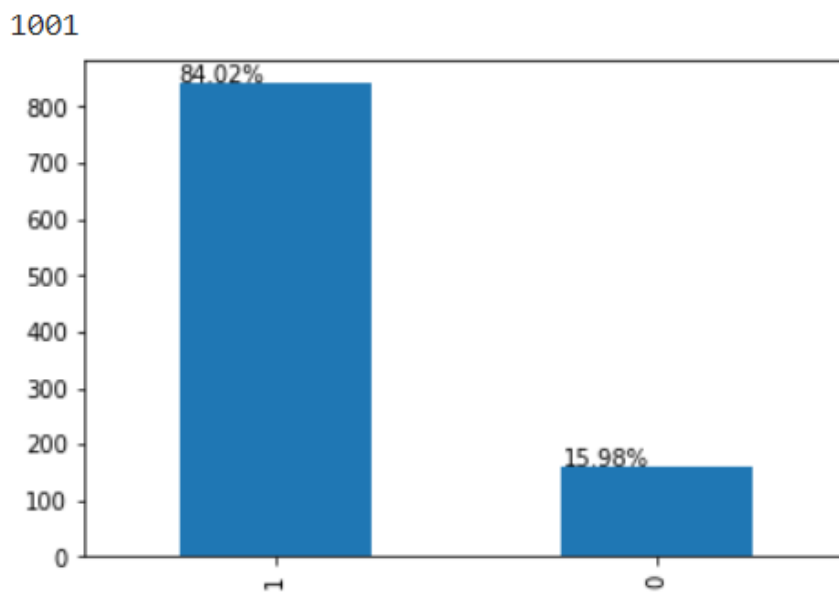
Tweet	Probabilitas	
	Positif	Negatif
Tweet 7	0,027	0,02439
Tweet 8	0,001095	0,00118976
Tweet 9	0,0000098415	0,000055296



Setelah melakukan pengujian probabilitas maka dapat dianalisis hasil dari data pengujian pertama, yaitu Tweet7, kita dapat menyimpulkan bahwa *tweet* tersebut masuk kedalam kategori positif karena nilai probabilitas positif yang dimiliki *tweet7* lebih besar dari pada nilai probabilitas negatif yang dimiliki *tweet7*. Pada *tweet 8* mengandung nilai positif lebih kecil dibanding probabilitas negatif maka dapat disimpulkan mengandung nilai negatif. Sedangkan pada *tweet 9* nilai *probabilitas* positif lebih kecil dibanding *probabilitas* negatif sehingga pada *tweet9* dinyatakan ke dalam kategori sentimen negatif. Hasil kategori dapat dilihat pada tabel 7.

Tabel 7. Hasil Kategori Data Testing

Tweet	Fitur	Kategori
Tweet7	Terimakasih telah meningkatkan perdagangan	Positif
Tweet8	Pemeliharaan mafia dilindungi pemerintah kemendag	Negatif
Tweet9	Kementerian yang membuat kecewa masyarakat dan sengsara	Negatif

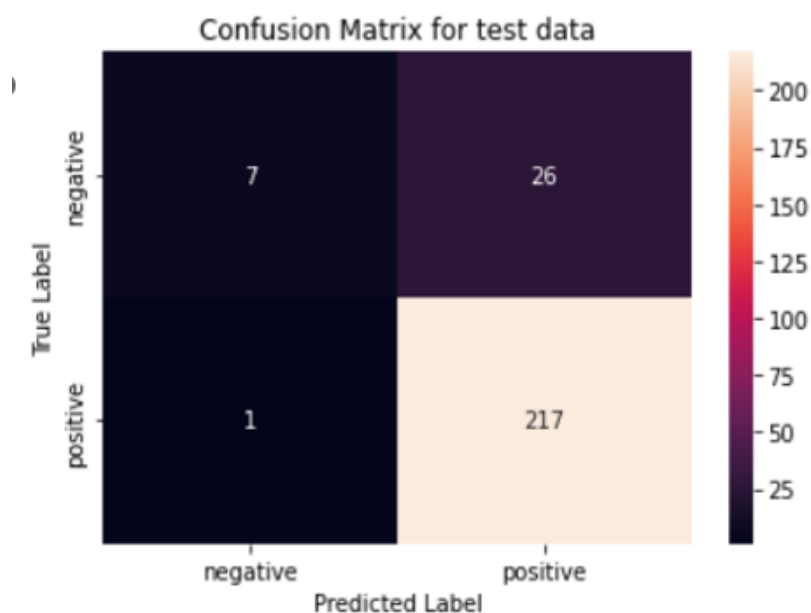


Gambar 4. Persentasi Tweet

Pada penelitian ini didapat hasil klasifikasi dari analisis sentimen dari 1000 data memiliki perbandingan 84,02% mengungkapkan nilai positif dengan ditandai dengan label 1 sedangkan sisanya 15,98% mengungkapkan pada nilai ungkapan negatif dengan ditandai label 0 seperti di tampilkan pada gambar 4.

### 3.4 Interpretation/evaluation

Pada tahap ini dilakukan evaluasi untuk mengukur validitas hasil klasifikasi dengan menghitung nilai *accuracy*, *pos pred value* dan *neg pred value*. Akurasi yang dihitung dari *confusion matrix* bertindak sebagai variabel evaluasi yang memiliki 4 nilai yang dijadikan acuan dalam perhitungannya yaitu *true positive rate* (TP rate), *true negative rate* (TN rate), *false positive rate* (FP rate) dan *false negative rate* (FN rate). Selanjutnya dilakukan penghitungan *Confusion Matrix* untuk mengetahui tingkat akurasi dari analisis yang dilakukan seperti pada gambar 5.



Gambar 5 Confusion Matrix

Berdasarkan gambar 5 dapat disimpulkan bahwa prediksi negatif dapat memprediksi dengan benar sebanyak 7 sedangkan kesalahan prediksi sebanyak 1 masuk kategori positif. Sedangkan hasil prediksi positif memprediksi dengan benar sebanyak 217 dan kesalahan prediksi sebanyak 26 masuk kedalam kategori negative seperti ditampilkan pada tabel 8.

Tabel 8. Prediksi Confusion Matrix

Data	Benar	Salah
Positif	217	26
Negatif	7	1

Berdasarkan tabel 8 dari hasil tabel prediksi diatas maka dapat kita ketahui tingkat akurasi yang didapat dengan menggunakan perhitungan manual sebagai berikut:

$$\begin{aligned}
 \text{Akurasi} &= \frac{\text{True Positive} + \text{True Negatie}}{\text{Total data yang diuji}} \times 100\% \\
 &= \frac{217 + 7}{7 + 26 + 1 + 217} \times 100\% \\
 &= 89,24\%
 \end{aligned}$$

Akurasi mewakili dari akurasi model yang dibuat dapat mengklasifikasikan data dengan benar. Akurasi diperoleh dengan menghitung rasio prediksi yang benar terhadap total data. Dengan mengetahui nilai akurasi kinerja seluruh sistem, maka dapat menentukan tingkat kemampuan sistem untuk menemukan akurasi antara informasi yang diinginkan pemakai dan jawaban yang diberikan sistem. Tingkat keberhasilan sistem yang menemukan informasi-akurasi pada penelitian ini adalah 89,24%.

Dari analisis ini wordcloud kumpulan kata-kata yang sering muncul pada kategori analisis sentimen yang mengandung nilai positif adalah ditunjukkan pada gambar 6.



Gambar 6 Wordcloud Positif

Pada gambar *wordcloud* negatif merupakan kumpulan kata-kata yang sering muncul pada klasifikasi analisis sentimen yang mengandung nilai negatif.



Gambar 7 Wordcloud Negatif

#### 4. KESIMPULAN DAN SARAN

Berdasarkan analisis sentimen yang dilakukan oleh peneliti, dapat disimpulkan bahwa kesimpulan pada penelitian ini antara lain:

1. Berdasarkan opini yang diberikan masyarakat terhadap Kementerian Perdagangan Republik Indonesia di media sosial *twitter* sebanyak 1000 *tweet* menunjukkan bahwa persentase analisis sentimen yang di dapatkan adalah 84,02 % mengungkapkan pendapat yang mengandung nilai positif dan 15,98% mengungkapkan kepada opini yang mengandung nilai negatif.
2. Pengujian model sentimen analisis *naïve bayes* dapat bekerja dengan baik. Hal ini dilakukan dengan melakukan penghitungan nilai akurasi dan *confusion matrix* mendapatkan 89,24 %.

#### DAFTAR PUSTAKA

- [1] E. Juliyanto and F. Rusdi, "Strategi Penyampaian Informasi Melalui Instagram Dengan Tampilan Infografis (di Kementerian Perdagangan RI)," *Prologia*, vol. 2, no. 2, p. 298, 2019, doi: 10.24912/pr.v2i2.3591.
- [2] M. Furqan, S. Mayang Sari, and P. Ilmu Komputer Fakultas Sains dan Teknologi, "Analisis Sentimen Menggunakan K-Nearest Neighbor Terhadap New Normal Masa Covid-19 Di Indonesia Sentiment Analysis using K-Nearest Neighbor towards the New Normal During the Covid-19 Period in Indonesia," *Techno.COM*, vol. 21, no. 1, pp. 52–61, 2022, [Online]. Available: [www.tripadvisor.com](http://www.tripadvisor.com).
- [3] J. Teknologi and I. Jtsi, "Analisis Sentimen Respon Masyarakat Terhadap Kabar Harian Covid-19 Pada Twitter Kementerian Kesehatan," vol. 2, no. 3, pp. 32–37, 2021.
- [4] F. F. Mailo and L. Lazuardi, "Analisis Sentimen Data Twitter Menggunakan Metode Text Mining Tentang Masalah Obesitas di Indonesia," *J. Inf. Syst. Public Heal.*, vol. 4, no. 1, pp. 28–36, 2019.
- [5] G. A. Buntoro, "Analisis Sentimen Calon Gubernur DKI Jakarta 2017 Di Twitter," *Integer J.*, vol. 2, no. 1, pp. 32–41, 2017.
- [6] S. P. Kristanto, J. A. Prasetyo, and E. Pramana, "Naive Bayes Classifier on Twitter Sentiment Analysis BPJS of HEALTH," *Proc. - 2019 2nd Int. Conf. Comput. Informatics Eng. Artif. Intell. Roles Ind. Revolut. 4.0, IC2IE 2019*, pp. 24–28, 2019, doi: 10.1109/IC2IE47452.2019.8940900.
- [7] K. Suppala and N. Rao, "Sentiment analysis using naïve bayes classifier," *Int. J. Innov. Technol. Explor. Eng.*, vol. 8, no. 8, pp. 264–269, 2019, doi: 10.14445/22312803/ijctt-v68i4p141.
- [8] A. Goel, J. Gautam, and S. Kumar, "Real time sentiment analysis of tweets using Naive Bayes," *Proc. 2016 2nd Int. Conf. Next Gener. Comput. Technol. NGCT 2016*, no. October, pp. 257–261, 2017, doi: 10.1109/NGCT.2016.7877424.
- [9] V. Malik and A. Kumar, "Analysis of Twitter Data Using Naive Bayes Algorithm," *Int. J. Recent Innov. Trends Comput. Commun.*, vol. 6, no. 4, pp. 120–125, 2018, [Online]. Available: <http://www.ijritcc.org>.
- [10] F. S. Mufidah *et al.*, "Analisis Sentimen Masyarakat terhadap Layanan Shopeefood Melalui Media Sosial Twitter dengan Algoritma Naive Bayes Classifier," *J. Inf. Syst.*, vol. 7, no. 1, pp. 14–25, 2022, doi: 10.33633/joins.v7i1.5883.
- [11] W. Wibowo and W. E. Novianti, "Analisis Sentimen Pengguna Twitter terhadap Program Kartu Prakerja di Tengah," *J. SAINS DAN SENI ITS*, vol. 11, no. 1, 2022.
- [12] D. A. Nugroho *et al.*, "Analisis Sentimen Data Presiden Jokowi Dengan Preprocessing Normalisasi Dan Stemming Menggunakan Metode Naive Bayes Dan SVM," *Semin. Nas. Teknol. Fak. Tek. Univ. Krisnadwipayana*, vol. 3, no. 1, pp. 1–11, 2021.
- [13] E. Nurhazidah, R. N. Ichsan, and S. Widiyanesti, "Analisis Sentimen Dan Jaringan Sosial Pada Penyebaran Informasi Vaksinasi Di Twitter," vol. 10, no. 1, pp. 24–35, 2022.

- [14] E. Indrayuni, “Klasifikasi Text Mining Review Produk Kosmetik Untuk Teks Bahasa Indonesia Menggunakan Algoritma Naive Bayes,” *J. Khatulistiwa Inform.*, vol. 7, no. 1, pp. 29–36, 2019, doi: 10.31294/jki.v7i1.1.
- [15] R. Gandhi, “Naive Bayes classifier,” vol. 6, pp. 1–9, 2018.
- [16] W. Priatna and R. Purnomo, “Comparison of Support Vector Machine and Artificial Neural Network Algorithm for Lecturer Performance Classification,” *Ijarcce*, vol. 10, no. 9, 2021, doi: 10.17148/ijarcce.2021.10901.
- [17] d n Katresna and f m Dzikry, “Implementasi Algoritma Naive Bayes Pada Klasifikasi Tweet Untuk Mengetahui Tingkat Kemalasan Siswa,” *J. Siliwangi Seri Sains Dan ...*, vol. 6, no. 2, pp. 66–70, 2020, [Online]. Available: <http://jurnal.unsil.ac.id/index.php/jssainstek/article/view/2528>.