

# Analisis Sentimen Ulasan Google Maps untuk Layanan Kesehatan Bandar Lampung menggunakan TF-IDF dan SVM

Pandra Insani Putra Azwar <sup>1</sup>, Putri Durrutol Shopia <sup>2</sup>, Catherine Firdhasari Maulina Sinaga <sup>3</sup>,

M. Syamsuddin Wisnubroto <sup>4</sup>, Fajri Farid <sup>5</sup>

<sup>1,2,3,4,5</sup>Program Studi Sains Data, Fakultas Sains, Institut Teknologi Sumatera, Lampung Selatan, Lampung, Indonesia

## Info Artikel

### Riwayat Artikel:

Received October 01, 2025

Revised October 03, 2025

Accepted February 13, 2026

**Abstract** – The development of digital technology has encouraged the public to utilize digital platforms like Google Maps as a primary medium for conveying experiences and opinions regarding public services, such as healthcare. These public reviews have become a crucial data source for assessing patient satisfaction and identifying service weaknesses. This study aims to analyze public sentiment towards healthcare services in Bandar Lampung City using machine learning. Support Vector Machine (SVM) was chosen as the text classification method, assisted by Term Frequency-Inverse Document Frequency (TF-IDF) weighting for word feature extraction. The research dataset was obtained through web scraping of 2,278 Google Maps reviews, which were then processed through preprocessing, data labeling, feature extraction, and model evaluation stages. The results of this study indicate a dominance of positive sentiment at 72.5% and negative sentiment at 27.5%. The SVM model with TF-IDF successfully achieved an accuracy of 89% with balanced precision, recall, and F1-score values, categorizing it as superior for text sentiment classification. These findings confirm that aspects of healthcare staff friendliness and service quality were positively rated, while administrative flows, drug distribution, and service speed remain areas of concern. This research utilizes Google Maps reviews as real-time data to evaluate healthcare services in Bandar Lampung City, a previously under-researched area. Sentiment analysis based on SVM and TF-IDF can serve as a data-driven evaluation tool to support strategic decision-making by the Bandar Lampung City Health Office in improving the quality of healthcare services.

**Keywords:** Sentiment Analysis, Google Maps, TF-IDF, SVM, Bandar Lampung City.

### \*Corresponding Author:

M. Syamsuddin Wisnubroto

Email:

syamsuddin.wisnubroto@sd.itea.ac.id



This is an open access article under the [CC BY 4.0](#) license.

**Abstrak** – Perkembangan teknologi digital mendorong masyarakat mulai memanfaatkan platform digital seperti Google Maps, sebagai media utama dalam menyampaikan pengalaman dan opini terhadap layanan publik, seperti layanan kesehatan. Ulasan publik ini menjadi sumber data penting untuk menilai kepuasan pasien sekaligus menilai kelemahan layanan. Penelitian ini bertujuan menganalisis sentimen masyarakat terhadap pelayanan kesehatan di Kota Bandar Lampung menggunakan machine learning. Support Vector Machine (SVM) dipilih sebagai metode klasifikasi teks dengan dibantu pembobotan Term Frequency-Inverse Document Frequency (TF-IDF) untuk mengekstraksi fitur kata. Dataset penelitian didapat melalui web scraping terhadap 2.278 ulasan Google Maps, yang kemudian diolah melalui tahapan preprocessing, pelabelan data, ekstraksi fitur, serta evaluasi model. Hasil penelitian ini menunjukkan dominasi sentimen positif sebesar 72,5% dan sentimen negatif sebesar 27,5%. Model SVM dengan TF-IDF berhasil mendapatkan akurasi 89% dengan nilai precision, recall, dan F1-score yang seimbang, sehingga dapat dikategorikan unggul untuk klasifikasi sentimen teks. Hasil penelitian ini menegaskan bahwa aspek keramahan tenaga kesehatan dan kualitas layanan dinilai positif, sementara alur administrasi, distribusi obat, dan kecepatan pelayanan masih menjadi perhatian. Penelitian menggunakan ulasan Google Maps sebagai data real-time untuk mengevaluasi pelayanan kesehatan di Kota Bandar Lampung, yang belum banyak diteliti sebelumnya. Analisis sentimen berbasis SVM dan TF-IDF dapat menjadi sarana evaluasi berbasis data yang mendukung pengambilan keputusan strategis Dinas Kesehatan Kota Bandar Lampung dalam meningkatkan mutu layanan kesehatan.

**Kata Kunci:** Analisis Sentimen, Google Maps, TF-IDF, SVM, Kota Bandar Lampung.

## I. PENDAHULUAN

Media sosial dan *platform daring* seperti Google Maps telah menjadi sarana vital di era digital bagi masyarakat untuk berbagi pengalaman mengenai layanan publik. Ulasan publik pada platform ini kini dimanfaatkan oleh penyedia layanan kesehatan sebagai instrumen transparansi untuk menilai kualitas pelayanan [1]. Di Kota Bandar Lampung, fasilitas kesehatan seperti puskesmas dan Rumah Sakit Umum Daerah (RSUD) memegang peran sentral dalam pelayanan masyarakat. Namun, kinerja fasilitas ini sering mendapat sorotan. Penelitian terdahulu menunjukkan adanya variasi kualitas layanan antar fasilitas di RSUD Abdoel Moeloek yang memiliki kelebihan fasilitas namun masih menghadapi tantangan efektivitas dibandingkan rumah sakit swasta, sementara kepuasan pasien di tingkat puskesmas sangat dipengaruhi oleh kualitas SDM dan manajemen fasilitas [2], [3]. Meskipun beberapa puskesmas

dinilai memuaskan, kendala struktural terkait keterbatasan kapasitas dan distribusi fasilitas yang belum merata masih menjadi keluhan utama yang memicu respons beragam dari publik [4].

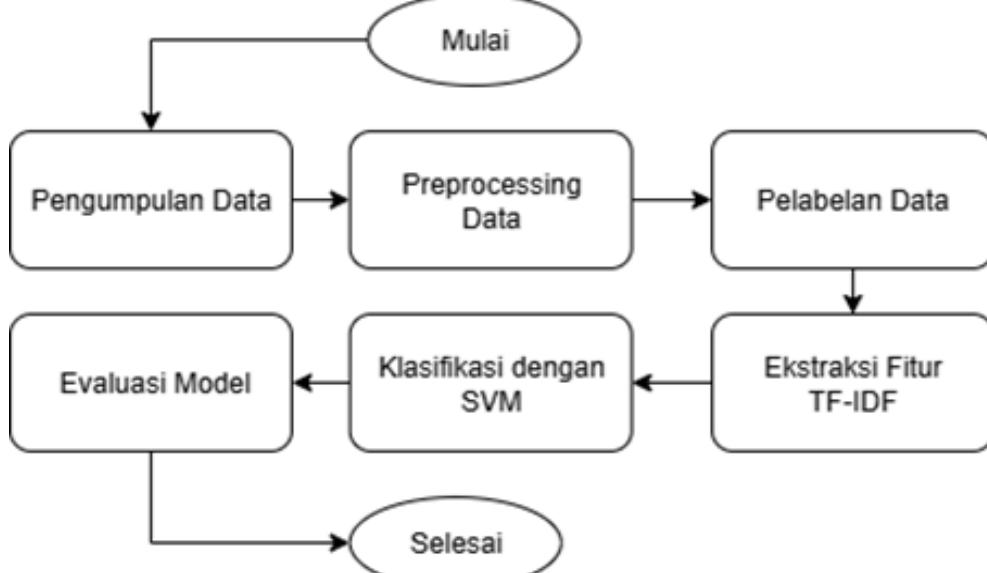
Analisis sentimen merupakan cabang dari *Natural Language Processing* (NLP) yang digunakan untuk mengumpulkan dan mengklasifikasikan pendapat masyarakat dari teks [5]. Metode ini efektif mengubah data teks mentah menjadi wawasan strategis. Berbagai penelitian telah membuktikan keunggulannya. Penggunaan algoritma BERT pada ulasan Google Maps, berhasil mengidentifikasi faktor determinan kepuasan pasien secara rinci [6]. Penelitian lain menunjukkan bahwa *Support Vector Machine* (SVM) dengan seleksi fitur *Chi-Square* mampu menghasilkan akurasi tinggi dalam klasifikasi opini [7]. Lebih lanjut, kombinasi SVM dengan pembobotan TF-IDF N-Gram terbukti meningkatkan kinerja model secara signifikan dibanding pembobotan kata sederhana [8].

*Support Vector Machine* (SVM) dan TF-IDF dinilai sangat efektif untuk tugas klasifikasi teks. SVM unggul dalam mengolah data berdimensi tinggi meskipun dengan sampel terbatas, sedangkan TF-IDF mampu memberikan bobot relevansi yang akurat pada kata-kata penting [7], [8]. Mayoritas penelitian sebelumnya berfokus pada layanan kesehatan secara umum atau menggunakan data media sosial seperti YouTube yang memiliki karakteristik berbeda [9]. Belum banyak penelitian yang secara spesifik menganalisis sentimen layanan kesehatan lokal di Kota Bandar Lampung menggunakan ulasan Google Maps. Keterbatasan ini menunjukkan adanya kebutuhan untuk sebuah penelitian yang berfokus pada konteks lokal untuk menghasilkan wawasan yang lebih relevan agar dapat ditindaklanjuti.

Tujuan dari penelitian ini adalah untuk menganalisis sentimen masyarakat terhadap pelayanan kesehatan di Kota Bandar Lampung dengan menggunakan kombinasi TF-IDF dan SVM. Hasil penelitian ini diharapkan dapat menjadi alat evaluasi strategis yang dapat digunakan oleh Dinas Kesehatan Kota Bandar Lampung, khususnya dalam aspek-aspek yang menjadi keluhan masyarakat seperti alur administrasi dan kecepatan pelayanan. Keunikan dalam penelitian ini adalah dalam pemanfaatan ulasan Google Maps sebagai gambaran opini publik yang *valid*, dengan fokus pada layanan kesehatan di Kota Bandar Lampung yang masih jarang diteliti, serta penggunaan kombinasi TF-IDF dan SVM pada data ulasan tersebut. Penelitian ini diharapkan dapat menyediakan landasan berbasis data dari opini digital yang relevan untuk mendukung peningkatan mutu layanan kesehatan di Kota Bandar Lampung.

## II. METODE

Metodologi yang digunakan dalam penelitian ini mengikuti alur kerja dari beberapa penelitian sebelumnya yang telah berhasil menerapkan SVM untuk menganalisis sentimen di bidang layanan kesehatan. Penelitian sebelumnya mengevaluasi opini masyarakat terhadap layanan BPJS telah membuktikan efektivitas alur kerja yang mencakup *pre-processing* data, pelabelan data, dan evaluasi model menggunakan metrik standar (*accuracy*, *precision*, *recall*, dan *F1-score*). Hasil penelitian sebelumnya menjalankan serangkaian tahapan metodologis yang sistematis, dimulai dari pengambilan data melalui *web scraping* ulasan di Google Maps menggunakan SerpAPI. Data mentah kemudian melalui tahap *pre-processing* teks yang komprehensif (*case folding*, *cleaning*, normalisasi, *tokenizing*, *stopword removal*, *stemming*). Tahap berikutnya adalah pelabelan data, ekstraksi fitur dengan pembobotan TF-IDF, klasifikasi menggunakan SVM, dan diakhiri dengan evaluasi performa model menggunakan *confusion matrix* [10].



Gambar 1. Alur penelitian

#### A. Alur Penelitian

Alur pada penelitian ini dibangun untuk analisis sentimen terhadap layanan kesehatan Kota Bandar Lampung menggunakan TF-IDF dan SVM meliputi beberapa langkah. Langkah-langkah tersebut antara lain, pengambilan data dengan teknik *web scraping* melalui SerpAPI, *pre-processing* data yang terdiri dari *case folding*, *cleaning*, normalisasi, *tokenizing*, *stopword removal*, dan *stemming*. Kemudian dilanjutkan pelabelan data, ekstraksi fitur menggunakan TF-IDF, klasifikasi dengan SVM, serta evaluasi model. Alur metode penelitian disajikan pada Gambar 1.

#### B. Data dan Sumber Data

Data penelitian ini dikumpulkan melalui teknik *web scraping* menggunakan layanan *Application Programming Interface* (API) SerpAPI pada platform Google Maps. Objek penelitian difokuskan pada ulasan masyarakat terhadap 15 Puskesmas dan 1 Rumah Sakit Umum Daerah (RSUD) di Kota Bandar Lampung. Pengambilan data dilakukan pada periode 14 Juli 2025 hingga 22 Agustus 2025. Total dataset diperoleh sebanyak 2.278 ulasan yang disimpan dalam format CSV untuk menjaga struktur data agar mudah diolah [11]. Penggunaan SerpAPI dipilih karena efisiensinya dalam mengekstraksi ulasan, rating, dan metadata tanpa perlu membangun *crawler* manual yang kompleks.

#### C. Pra-pemrosesan Data

Tahap *pre-processing* dalam pengolahan data digunakan untuk mengubah data mentah menjadi lebih terstruktur dan siap dianalisis. Dalam penelitian analisis sentimen ini, tahap *pre-processing* dilakukan melalui serangkaian proses meliputi:

- Case Folding: Mengubah seluruh karakter dalam teks menjadi huruf kecil menggunakan pustaka NLTK. Langkah ini bertujuan untuk menyeragamkan format data sehingga proses analisis dapat dilakukan secara konsisten.
- Cleaning: Menghapus nilai yang hilang (missing value) serta mengeliminasi elemen-elemen yang tidak relevan, seperti tanda baca, simbol, maupun karakter yang tidak sesuai dengan kaidah penulisan menggunakan pustaka regular expression (re).
- Normalisasi: Tahap mengubah kata-kata tidak baku atau berbentuk singkatan menjadi bentuk kata yang standar dan umum sesuai dengan rujukan kamus yang digunakan.
- Tokenizing: Proses memecah teks atau kalimat utuh menjadi unit-unit kata yang terstruktur. Sebagai contoh, kalimat "buruk banget awal dari" akan diubah menjadi token kata "buruk", "banget", "awal", dan "dari".
- Stopword Removal: Tahap dalam pemrosesan teks yang berfungsi menghilangkan kata-kata umum yang dianggap tidak memiliki kontribusi signifikan terhadap makna kalimat.
- Stemming: Proses menghilangkan bentuk infleksi pada kata, yaitu penambahan awalan atau akhiran yang membentuk kata turunan. Dengan demikian, kata yang digunakan dalam analisis bukan lagi kata berimbahan, melainkan akar katanya.

#### D. Pelabelan Data

Penelitian sebelumnya menunjukkan tren penggunaan model *deep learning* berbasis *transformer*, khususnya RoBERTa (*Robustly Optimized BERT Pretraining Approach*), untuk pelabelan data otomatis pada analisis sentimen teks berbahasa Indonesia. Model ini menjadi pilihan utama karena kemampuannya memahami konteks kalimat dan menghasilkan klasifikasi sentimen yang akurat, menggantikan proses pelabelan manual yang subjektif dan memakan waktu lama [12]. Keunggulan utama metode ini adalah efisiensi dan konsistensi, di mana pelabelan ribuan data dapat diselesaikan dalam hitungan menit. Penelitian terdahulu menggunakan model ini untuk melabeli *tweet* terkait profesi PNS ke dalam kategori positif, negatif, dan netral [13]. Metode serupa juga diterapkan pada analisis sentimen ulasan game Mobile Legends di Google Play Store dan Youtube [12]. Meskipun akurasi model bervariasi tergantung konteks dataset, pendekatan ini terbukti efektif mengatasi kesulitan dalam memperoleh data berlabel [14].

Penelitian ini menerapkan pelabelan data menggunakan model pretrained *wlwo/indonesian-roberta-base-sentiment-classifier* yang telah dilatih pada korpus Bahasa Indonesia. Model ini dirancang mengklasifikasikan teks ke dalam tiga kategori: positif, negatif, dan netral. Namun, karena penelitian ini berfokus pada klasifikasi biner, ulasan dengan prediksi sentimen netral dieliminasi dari dataset. Pendekatan ini berpotensi menghasilkan distribusi kelas yang tidak seimbang (*imbalanced*). Oleh karena itu, untuk mengatasi potensi bias evaluasi akibat dominasi satu kelas, pengukuran performa model tidak hanya akan mengandalkan akurasi, tetapi juga menggunakan metrik *F1-Score* untuk memberikan penilaian yang lebih menyeluruh.

#### E. Ekstraksi Fitur dengan TF-IDF

Setelah data terlabeli, dilakukan pembobotan kata menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF). Metode ini memberikan bobot tinggi pada kata yang jarang muncul secara global namun sering

muncul dalam dokumen tertentu, yang mengindikasikan relevansi kata tersebut [15]. Bobot kata dihitung menggunakan Persamaan 1 dan 2.

$$Wt, d = TFt, d \times IDFt \quad (1)$$

$$IDFt = \log \log \left( \frac{N}{NT} \right) \quad (2)$$

Keterangan :

$Wt, d$	: Bobot dari t ( <i>term</i> ) dalam suatu dokumen
$TFt, d$	: Frekuensi kemunculan t ( <i>term</i> ) dalam dokumen d
$IDFt$	: Inverse Document Frequency
$N$	: Jumlah dokumen
$NT$	: Jumlah dokumen yang mengandung <i>term</i>

#### F. Klasifikasi dengan Support Vector Machine (SVM)

Algoritma SVM digunakan untuk membangun model klasifikasi sentimen. Prinsip dasar SVM adalah menemukan *hyperplane* terbaik yang memisahkan dua kelas data dengan *margin* maksimal dalam ruang vektor berdimensi tinggi [16]. Penelitian ini menggunakan fungsi *Kernel Linear* karena data teks memiliki fitur yang sangat banyak dan tersebar (*sparse*), sehingga pemisahan linear seringkali memberikan performa terbaik dan komputasi yang efisien [17]. Fungsi keputusan untuk kernel linear didefinisikan pada Persamaan 3.

$$K(x_i, x) = x_i^T x \quad (3)$$

Keterangan :

$x_i$	: Data latih
$x$	: Data Uji

#### G. Evaluasi Model

Performa model dievaluasi menggunakan *confusion matrix* untuk mengukur seberapa tepat prediksi model terhadap data actual yang terbagi menjadi empat, dijelaskan sebagai berikut:

- True Positives (TP): Jumlah data kelas positif yang tepat diprediksi sebagai kelas positif.
- True Negatives (TN): Jumlah data kelas negatif yang tepat diprediksi sebagai kelas negatif.
- False Positives (FP): Jumlah data kelas negatif yang salah diprediksi menjadi kelas positif.
- False Negative (FN): Jumlah data kelas positif yang salah diprediksi menjadi kelas negatif.

Metrik yang digunakan dalam penelitian ini meliputi *Accuracy*, *Precision*, *Recall*, dan *F1-Score* [8]. Penggunaan *F1-Score* menjadi krusial dalam penelitian ini untuk menangani potensi ketidakseimbangan kelas data (*imbalanced data*) hasil pelabelan otomatis. Perhitungan metrik dilakukan berdasarkan Persamaan 4-7.

$$Precision = \frac{TP}{TP + FP} \quad (4)$$

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (6)$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (7)$$

### III. HASIL DAN PEMBAHASAN

Penelitian ini menggunakan dataset yang berisi 2,278 ulasan masyarakat tentang pelayanan kesehatan yang ada di Kota Bandar Lampung melalui Google Maps. *Pre-processing* dilakukan beberapa langkah, yaitu *case folding*, *cleaning*, normalisasi, *tokenizing*, *stopword removal*, dan *stemming* yang menghasilkan data 1,912. Setelah dilakukan *pre-processing* dilakukan pelabelan data menggunakan model *pretrained transformer Indonesian RoBERTa Base Sentiment Classifier* yang telah dilatih pada korpus Bahasa Indonesia untuk mengklasifikasikan setiap teks menjadi positif dan negatif. Setelah pelabelan data dilakukan ekstraksi fitur dengan TF-IDF. Hasil perubahan data sebelum *pre-processing* dan setelah *pre-processing* ditunjukkan pada Tabel 1.

TABEL 1  
HASIL PRE-PROCESSING

Ulasan	Preprocessing	Hasil
Buruk banget awal dari mendahulukan orang yang dikenal adalah salah satu cara orang tidak profesional banget dalam dunia kerja, bagaimana Indonesia akan maju sedangkan orang yang seperti ini masih banyak banget haduh haduhhhh buruk!	Case folding	buruk banget awal dari mendahulukan orang yang dikenal adalah salah satu cara orang tidak profesional banget dalam dunia kerja, bagaimana indonesia akan maju sedangkan orang yang seperti ini masih banyak banget haduh haduhhhh buruk!
	Cleaning	buruk banget awal dari mendahulukan orang yang dikenal adalah salah satu cara orang tidak profesional banget dalam dunia kerja bagaimana indonesia akan maju sedangkan orang yang seperti ini masih banyak banget haduh haduhhhh buruk
	Normalisasi	buruk banget awal dari mendahulukan orang yang dikenal adalah salah satu cara orang tidak profesional banget dalam dunia kerja bagaimana indonesia akan maju sedangkan orang yang seperti ini masih banyak banget haduh haduhhhh buruk
	Tokenizing	"buruk", "banget", "awal", "dari", "mendahulukan", "orang", "yang", "di", "kenal", "adalah", "salah", "satu", "cara", "orang", "tidak", "profesional", "banget", "dalam", "dunia", "kerja", "bagaimana", "indonesia", "akan", "maju", "sedangkan", "orang", "yang", "seperti", "ini", "masih", "banyak", "banget", "haduh", "haduhhhh", "buruk"
	Stopword removal	"buruk", "banget", "mendahulukan", "orang", "kenal", "salah", "orang", "professional", "banget", "dunia", "kerja", "Indonesia", "maju", "orang", "banget", "haduh", "haduhhhh", "buruk"
	Stemming	buruk banget dahulu orang kenal salah orang profesional banget dunia kerja indonesia maju orang banget haduh haduhhhh buruk

Proses pelabelan data dalam penelitian ini dilakukan secara otomatis dengan menggunakan model *pre-trained* berbasis *transformer*, yaitu *RoBERTa Base Sentiment Classifier* yang tersedia melalui *platform Hugging Face*. Model tersebut telah dilatih khusus untuk tugas klasifikasi sentimen berbahasa Indonesia, sehingga relevan dengan konteks penelitian ini. Tahap berikutnya adalah memberi setiap ulasan label sentimen positif, negatif, maupun netral, sesuai hasil prediksi model, sebagaimana ditunjukkan pada Tabel 2.

TABEL 2  
HASIL PELABELAN DATA

Positif	Negatif	Netral
1387	525	333

Proses pelabelan pada penelitian ini difokuskan pada dua polaritas sentimen utama yaitu positif dan negatif, ulasan dengan label sentimen netral dieliminasi dari dataset, karena tidak merepresentasikan polaritas sentimen yang menjadi fokus utama pada penelitian. Hasil dari eliminasi data netral ditampilkan pada Tabel 3.

TABEL 3  
HASIL PENGHAPUSAN DATA NETRAL

Positif	Negatif
1387	525

Setelah seluruh tahap *pre-processing* selesai, langkah selanjutnya adalah melakukan pembobotan fitur untuk mengukur relevansi setiap *term* (kata) dalam ulasan. Tidak semua *term* (kata) memiliki bobot TF-IDF yang sama. Kata yang sering muncul seperti "dan" atau "yang" (meskipun sebagian besar sudah dihilangkan di tahapan *stopword removal*) memiliki nilai bobot yang rendah. Oleh karena itu, penelitian ini menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF) untuk menilai seberapa relevan sebuah kata dalam sebuah dokumen (ulasan) terhadap keseluruhan dokumen. Sepuluh kata dengan bobot rata-rata TF-IDF tertinggi ditampilkan pada Tabel 4 yang menunjukkan bahwa istilah "layan" memiliki bobot tertinggi (0,130424) dalam ulasan layanan kesehatan Kota Bandar Lampung, menunjukkan bahwa pelayanan adalah elemen paling penting dan paling sering dibahas. *Term* (kata) lain yang dominan, seperti "ramah" (0,083769), "bagus" (0,051017), dan "cepat" (0,044083), secara keseluruhan, pandangan publik terpusat pada penilaian pada kualitas pelayanan dan efisiensi staf medis.

TABEL 4

PEMBOBOTAN KATA DENGAN TF-IDF

Peringkat	Ulasan	Bobot rata-rata TF-IDF
1	layan	0,130424
2	ramah	0,083769
3	bagus	0,051017
4	cepat	0,044083
5	bersih	0,026107
6	ruang	0,025493
7	obat	0,021688
8	muas	0,021370
9	banget	0,018634
10	nyaman	0,017259



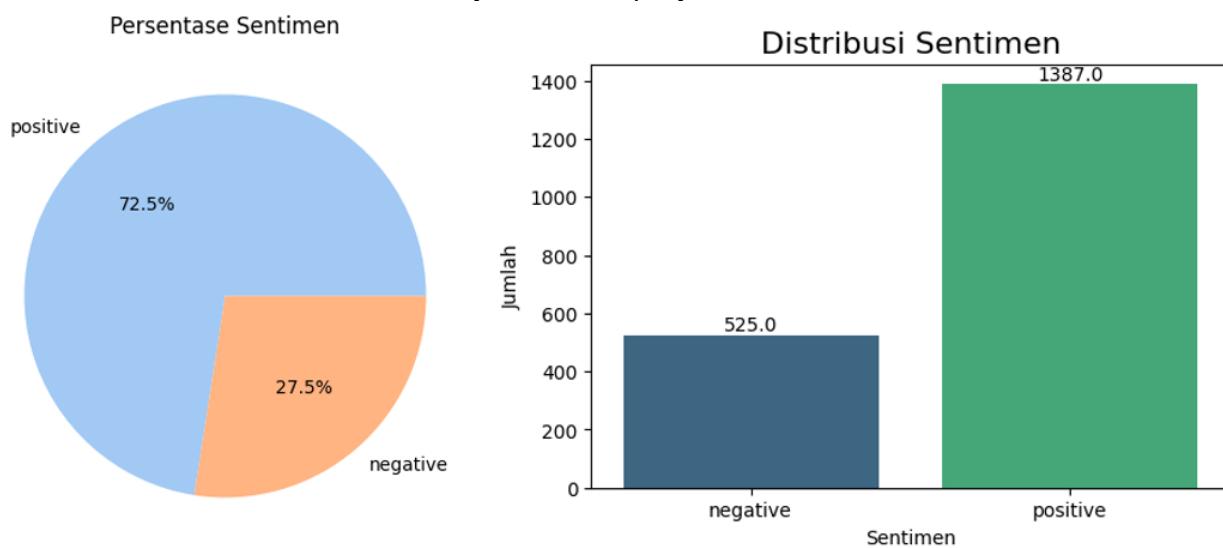
Gambar 2. Visualisasi word cloud

*Word Cloud* pada Gambar 2 menampilkan distribusi kata yang paling sering muncul dalam ulasan masyarakat terkait pelayanan kesehatan di Kota Bandar Lampung. Semakin besar ukuran sebuah kata, semakin tinggi frekuensi kemunculannya dalam *dataset*. Dari hasil visualisasi terlihat bahwa kata “pelayanan”, “bagus”, “berobat”, “ramah”, dan “bersih” mendominasi kumpulan ulasan, yang menunjukkan fokus masyarakat pada kualitas pelayanan, interaksi dengan tenaga medis, serta pengalaman selama mendapatkan layanan kesehatan. Visualisasi ini membantu memberikan gambaran awal mengenai aspek layanan yang menjadi perhatian utama pasien dan pengunjung layanan kesehatan yang ada di Kota Bandar Lampung, sebelum dilakukan analisis sentimen lebih mendalam dengan metode *Support Vector Machine* (SVM).

**TABEL 5**  
**HASIL EVALUASI CONFUSION MATRIX**

Actual	Predicted	
	Positive	Negative
Positive	247	18
Negative	24	94
Accuracy	0,89	
Precision	0,91	0,84
Recall	0,93	0,80
FI-Score	0,92	0,82

Berdasarkan Tabel 5, model SVM dengan TF-IDF diuji pada tugas klasifikasi biner (positif dan negatif) setelah ulasan berlabel netral dieliminasi dari *dataset* (sesuai tahap metodologi sebelumnya). Model berhasil mengklasifikasikan 247 ulasan positif dan 94 ulasan negatif secara benar, dengan kesalahan 18 ulasan positif dan 24 ulasan negatif. Model mencapai *accuracy* 89% dengan *precision* 0,91(positif) dan 0,84 (negatif), *recall* 0,93 (positif) dan 0,82 (negatif), serta *F1-score* 0,92 (positif) dan 0,82 (negatif). Hasil ini menunjukkan bahwa metode SVM cukup baik untuk analisis sentimen biner ulasan masyarakat terkait pelayanan kesehatan.



Gambar 3. Diagram hasil distribusi label sentimen

Gambar 3 menampilkan hasil dari pelabelan menggunakan model *RoBERTa Base Sentiment Classifier* terhadap 1.976 data yang *valid*, diperoleh hasil distribusi sentimen sebagai berikut:

- Sentimen Positif: 1387 data (72,5%).
- Sentimen Negatif: 525 data (27,5%).

Distribusi sentimen menampilkan dominasi sentimen positif (72,5%) yang menunjukkan persepsi masyarakat cukup baik terhadap pelayanan kesehatan di Kota Bandar Lampung. Proporsi sentimen negatif sebesar (27,5%) mengindikasikan bahwa ada yang perlu diperbaiki dari aspek pelayanan di Kota Bandar Lampung.

Hasil penelitian menunjukkan bahwa mayoritas ulasan masyarakat terhadap pelayanan kesehatan di Kota Bandar Lampung memiliki sentimen positif. Dari 1,912 data yang dianalisis, sebanyak 72,5% tergolong sentimen positif, sedangkan 27,5% tergolong sentimen negatif. Hasil ini menggambarkan bahwa mayoritas masyarakat memberikan apresiasi terhadap pelayanan kesehatan. Persentase sentimen negatif yang cukup tinggi tidak dapat diabaikan. Keluhan masyarakat banyak berkaitan dengan kecepatan layanan, terutama dalam proses antrian, pengambilan obat di apotek, serta prosedur administratif seperti penggunaan BPJS. Kata-kata seperti “judes”, “berobat”, dan “pelayanan” sering muncul dalam ulasan negatif, yang menandakan adanya kendala teknis dan efisiensi pelayanan. Dengan *accuracy* mencapai 89% dan nilai *accuracy*, *precision*, *recall*, dan *F1-score* yang seimbang, evaluasi model menggunakan *Support Vector Machine* (SVM) dengan ekstraksi fitur TF-IDF menunjukkan bahwa model ini dapat diandalkan untuk mengklasifikasikan ulasan masyarakat secara konsisten, baik positif maupun negatif.

Hasil penelitian ini memberikan implikasi praktis yang berarti analisis sentimen dapat digunakan oleh Dinas Kesehatan Kota Bandar Lampung untuk membuat keputusan tentang bagaimana meningkatkan pelayanan kesehatan. Sentimen positif menunjukkan aspek keberhasilan tertentu, seperti sikap ramah tenaga medis, yang harus dipertahankan dan diperkuat. Sementara itu, sentimen negatif menunjukkan hal-hal yang perlu segera diperbaiki, seperti kecepatan pelayanan dan kelancaran distribusi obat. Penelitian ini tidak hanya menghasilkan model klasifikasi sentimen yang efektif, tetapi juga menyediakan informasi berbasis data yang dapat mendukung pengambilan keputusan strategis dalam rangka meningkatkan kualitas pelayanan kesehatan di Kota Bandar Lampung.

#### IV. KESIMPULAN

Hasil penelitian ini menunjukkan bahwa sentimen masyarakat terhadap pelayanan kesehatan di Kota Bandar Lampung lebih unggul oleh ulasan positif (72,5%) dibandingkan ulasan negatif (27,5%). Model klasifikasi *Support Vector Machine* (SVM) dengan ekstraksi fitur pembobotan TF-IDF terbukti andal dengan tingkat akurasi mencapai 89%. Berdasarkan hasil temuan ini, direkomendasikan agar Dinas Kesehatan Kota Bandar Lampung mempertahankan aspek-aspek yang dinilai positif oleh masyarakat, seperti keramahan tenaga medis. Meskipun begitu, area yang menjadi sumber keluhan utama khususnya terkait kecepatan pelayanan, distribusi obat, dan alur administrasi memerlukan perhatian dan perbaikan segera. Penelitian ini menegaskan bahwa analisis sentimen dapat berfungsi sebagai sarana evaluasi yang valid untuk merumuskan kebijakan berbasis data yang berpusat pada kebutuhan pasien.

#### UCAPAN TERIMA KASIH

Terima kasih kepada Dosen Program Studi Sains Data, Institut Teknologi Sumatera atas bimbingan dan masukan berharga selama penelitian ini. Ucapan terima kasih juga ditujukan kepada keluarga, teman, dan semua pihak yang telah memberikan dukungan, baik secara moral maupun material. Semua bantuan yang diberikan sangat berarti dalam menyelesaikan penelitian ini.

#### DAFTAR PUSTAKA

- [1] Sidabutar, T. G. W. M., & Juardi, D. (2025). Analisis sentimen masyarakat terhadap penggunaan Halodoc sebagai layanan telemedicine di Indonesia. *Jurnal Informatika dan Teknik Elektro Terapan*, 13(1). <https://doi.org/10.23960/jitet.v13i1.5682>
- [2] Efriyady, D. (2023). Kepuasaan Pasien pada Puskesmas Way Halim Bandar Lampung. *Journal of Management and Creative Business*, 1(4), 425-439. <https://doi.org/10.30640/jmcbus.v1i4.1722>
- [3] HERIANSA, N. (2024). Analisis Tingkat Kepuasan Pasien Metode Servqual terhadap Mutu Pelayanan Laboratorium di UPT Puskesmas Rawat Inap Way Kandis Bandar Lampung (Doctoral dissertation, POLTEKKES KEMENKES TANJUNGPINANG).
- [4] Lesmana, A. V. D., Safira , H., & Lukman , M. (2025). Perbandingan Pelayanan Publik Antara Rumah Sakit Swasta dan Rumah Sakit Negeri di Lampung: Studi Kasus RSUD Abdoel Moeloek dan Rumah Sakit Advent Bandar Lampung. *Komsospol*, 5(1), 21–27. <https://doi.org/10.47637/komsospol.v5i1.1569>
- [5] Angel, A. C. T., Pranatwijaya, V. H., & Widiatry, W. (2024). Analisis sentimen dan emosi dari ulasan Google Maps untuk layanan rumah sakit di Palangka Raya menggunakan machine learning. *KONSTELASI: Konvergensi Teknologi dan Sistem Informasi*, 4(1), 35–49. <https://doi.org/10.24002/konstelasi.v4i1.8924>
- [6] Widagdo, A. S., Qodri, K. N., & Saputro, F. E. N. (2023). Analisis sentimen terhadap pelayanan Kesehatan berdasarkan ulasan Google Maps menggunakan BERT. *JURNAL FASILKOM*, 13(02), 326-333.
- [7] Hokijuliandy, E., Napitupulu, H., & Firdaniza, F. (2023). Analisis sentimen menggunakan metode klasifikasi support vector machine (SVM) dan seleksi fitur Chi-Square. *SisInfo: Jurnal Sistem Informasi dan Informatika*, 5(2), 40–49. <https://doi.org/10.37278/sisinfo.v5i2.670>

- [8] Arifin, N., Enri, U., & Sulistiyowati, N. (2021). Penerapan algoritma support vector machine (SVM) dengan TF-IDF N-Gram untuk text classification. STRING (Satuan Tulisan Riset dan Inovasi Teknologi), 6(2), 129–138. <https://doi.org/10.30998/string.v6i2.10133>
- [9] Sari, R. A. P., Kacung, S., & Santoso, B. (2025). ANALISIS SENTIMEN LAYANAN KESEHATAN BPJS MENGGUNAKAN METODE SVM. Jurnal Informatika Teknologi dan Sains (Jinteks), 7(2), 878-885. <https://doi.org/10.51401/jinteks.v7i2.5797>
- [10] Darwis, D., Pratiwi, E. S., & Pasaribu, A. F. O. (2020). Penerapan Algoritma SVM Untuk Analisis Sentimen Pada Data Twitter Komisi Pemberantasan Korupsi Republik Indonesia. Jurnal Ilmiah Edutic: Pendidikan dan Informatika, 7(1), 1-11. <https://doi.org/10.21107/edutic.v8i1.8779>
- [11] Gifari, O. I., Adha, M., Hendrawan, I. R., & Durrand, F. F. S. (2022). Analisis Sentimen Review Film Menggunakan TF-IDF dan Support Vector Machine. Journal of Information Technology, 2(1), 36-40. <https://doi.org/10.46229/jifotech.v2i1.330>
- [12] Pratama, M. R., Handayani, M. R., Yuniarti, W. D., & Umam, K. (2025). Model prediksi sentimen ulasan mobile legends di Google Play Store dan YouTube berbasis pelabelan otomatis roberta dan klasifikasi random forest. JSil Jurnal Sistem Informasi, 12(2), 39–45. <https://doi.org/10.30656/jsii.v12i2.10459>
- [13] Jaya, A. (2023). Analisis Sentimen Pandangan Public Profesi PNS (Pegawai Negeri Sipil) dari Twiter menerapkan indonesian Roberta Base Sentiment Classifier. Indonesian Journal of Data and Science, 4(1), 38-44. <https://doi.org/10.56705/ijdas.v4i1.66>
- [14] Kumar, A., & Wang, J. (2025). Numerical modelling in assisting machine learning for the enhanced condition monitoring of rotors. Results in Engineering, 107250. <https://doi.org/10.1016/j.rineng.2025.107250>
- [15] Yolanda, A. M., & Mulya, R. T. (2024). Implementasi Metode Support Vector Machine untuk Analisis Sentimen pada Ulasan Aplikasi Sayurbox di Google Play Store. VARIANSI: Journal of Statistics and Its application on Teaching and Research, 6(02), 76-83. [10.35580/variantsium258](https://doi.org/10.35580/variantsium258)
- [16] HAEKAL, M., & WIBOWO, W. C. (2023). Prediksi kualitas air sungai menggunakan metode pembelajaran mesin: Studi kasus Sungai Ciliwung: Prediction of river water quality using machine learning methods: Ciliwung River case study. Jurnal Teknologi Lingkungan, 24(2), 273-282. <https://doi.org/10.55981/jtl.2023.795>
- [17] Restiani, Y., & Purwadi, J. (2024). Support Vector Machine for Classification: A Mathematical and Scientific Approach in Data Analysis. Jurnal Penelitian Pendidikan IPA, 10(11), 9896-9903. <https://doi.org/10.29303/jppipa.v10i11.8122>