



Implementasi Text Mining dan Support Vector Machine untuk Analisis Sentimen Pelayanan Puskesmas

Sami Fitri^{1*}, Rizaldy Khair²

^{1*,2}Program Studi Sistem Informasi, Fakultas Ilmu Komputer dan Teknologi Informasi, Universitas Muhammadiyah Sumatera Utara, Medan, Indonesia
Email: ^{1*}fitriisami428@gmail.com, ²rizaldykhair@umsu.ac.id

Abstrak

Puskesmas sebagai fasilitas kesehatan tingkat pertama menghadapi keterbatasan dalam mengolah opini masyarakat yang tersebar di media sosial dan kanal daring secara sistematis dan real-time. Penelitian ini bertujuan mengembangkan sistem analisis sentimen berbasis *text mining* dan algoritma *Support Vector Machine* (SVM) untuk mengklasifikasikan opini masyarakat terhadap pelayanan Puskesmas Gurgur Pardomuan, Kabupaten Aceh Tenggara, ke dalam tiga kategori: positif, negatif, dan netral, sekaligus mengimplementasikannya dalam sistem berbasis web sebagai alat bantu evaluasi pelayanan. Penelitian ini menggunakan pendekatan kuantitatif yang dikombinasikan dengan pendekatan rekayasa sistem (*system engineering approach*). Data primer diperoleh melalui kuesioner daring dan ulasan masyarakat di media sosial, kemudian diproses melalui empat tahap *text preprocessing* (*case folding, tokenizing, stopword removal, stemming*), diekstraksi fitur menggunakan pembobotan TF-IDF (*Term Frequency-Inverse Document Frequency*), dan diklasifikasikan dengan SVM kernel linear berstrategikan *One-vs-Rest* (OvR). Model dilatih pada 200 data berlabel dengan komposisi 70 data positif (35%), 68 negatif (34%), dan 62 netral (31%), serta dievaluasi menggunakan *confusion matrix* pada 10 data uji independen. Evaluasi menunjukkan akurasi keseluruhan 70%, dengan performa sempurna pada kelas negatif (*precision* 100%, *recall* 100%, *F1-score* 100%), namun kelas netral masih terbatas (*precision* 50%, *recall* 33,33%, *F1-score* 40%) akibat ketimpangan jumlah data latih antarkelas. Pada konfigurasi pengujian dengan 30 data latih dan kernel linear, model mencapai akurasi 90%, *precision* 89,50%, *recall* 90,20%, dan *F1-score* 89,80%. Seluruh fitur sistem — input komentar, *training* model, klasifikasi sentimen, dan ekspor laporan — berhasil divalidasi melalui *black-box testing*. Temuan penelitian ini menunjukkan bahwa pendekatan *text mining* berbasis SVM dapat menjadi alternatif yang lebih responsif dibandingkan survei manual untuk memantau persepsi masyarakat terhadap layanan kesehatan primer di wilayah pedesaan Indonesia.

Kata Kunci: Analisis Sentimen, Puskesmas, Support Vector Machine, Text Mining, TF-IDF.

PENDAHULUAN

Transformasi digital telah mengubah cara masyarakat mengekspresikan kepuasan maupun ketidakpuasan terhadap layanan publik. Di bidang kesehatan, opini masyarakat yang sebelumnya hanya tersalurkan melalui survei atau kotak saran kini tersebar secara masif melalui media sosial, kuesioner daring, dan berbagai platform digital lainnya. Kondisi ini menciptakan peluang sekaligus tantangan bagi pengelola fasilitas kesehatan: di satu sisi, tersedia data umpan balik masyarakat yang kaya dan berkelanjutan; di sisi lain, volume dan heterogenitas data tersebut membuat analisis manual menjadi tidak praktis, rawan subjektivitas, dan lambat dalam menghasilkan informasi yang dapat ditindaklanjuti (Wankhade et al., 2022).

Puskesmas sebagai fasilitas kesehatan tingkat pertama (*primary health care*) memegang peran strategis dalam sistem kesehatan nasional Indonesia. Sebagaimana diamanatkan oleh Peraturan Menteri Kesehatan Nomor 43 Tahun 2019, puskesmas menyelenggarakan pelayanan promotif, preventif, kuratif, dan rehabilitatif untuk meningkatkan derajat kesehatan masyarakat di wilayah kerjanya. Keberhasilan puskesmas sangat dipengaruhi oleh kualitas layanan yang diberikan, yang pada gilirannya bergantung pada kemampuan institusi tersebut dalam merespons kebutuhan dan harapan masyarakat secara berkelanjutan (Murtini & Puspaningtyas, 2023). Namun, mekanisme evaluasi yang selama ini lazim digunakan yakni Survei Kepuasan Masyarakat (SKM) berbasis kuesioner tertutup memiliki keterbatasan mendasar: instrumen ini tidak mampu menangkap nuansa opini terbuka masyarakat, dilaksanakan periodik sehingga

tidak *real-time*, dan sering kali menghasilkan data yang terlalu agregat untuk memandu perbaikan layanan pada aspek spesifik.

Puskesmas Gurgur Pardomuan di Kabupaten Aceh Tenggara menghadapi persoalan tipikal puskesmas di wilayah pedesaan: data kepuasan masyarakat yang tercatat (65% berdasarkan survei Dinas Kesehatan Aceh Tenggara, 2022) masih jauh dari memuaskan, dengan keluhan utama pada waktu tunggu, kualitas layanan, dan fasilitas. Secara nasional, Kementerian Kesehatan RI (2023) mencatat bahwa sekitar 40% keluhan terhadap layanan kesehatan berasal dari layanan primer, diperparah oleh keterbatasan sumber daya. Namun demikian, puskesmas ini belum memiliki mekanisme untuk mengolah opini masyarakat yang beredar di media sosial secara sistematis, sehingga respons terhadap keluhan cenderung lambat dan reaktif.

Analisis sentimen (*sentiment analysis*) berbasis *text mining* dan *machine learning* telah terbukti mampu menjembatani kesenjangan tersebut. Teknik ini memungkinkan pengolahan data opini teks dalam jumlah besar secara otomatis, objektif, dan berkelanjutan (Wankhade et al., 2022). Di antara berbagai algoritma yang tersedia, *Support Vector Machine* (SVM) secara konsisten menunjukkan performa kompetitif pada tugas klasifikasi teks berdimensi tinggi karena kemampuannya memaksimalkan *margin* pemisah antarkelas (*hyperplane*) bahkan pada ruang fitur yang sangat besar (Guido et al., 2024). Kombinasi SVM dengan pembobotan TF-IDF sebagai ekstraksi fitur teks telah banyak diterapkan pada berbagai domain analisis sentimen, termasuk layanan kesehatan, dan menghasilkan performa yang andal (Khan et al., 2024; Imanuddin et al., 2023).

Namun, sebagian besar penelitian analisis sentimen berbasis SVM di ranah kesehatan Indonesia masih berfokus pada aplikasi kesehatan digital skala nasional (Imanuddin et al., 2023) atau ulasan platform e-health besar, dan belum menyoroti layanan puskesmas di wilayah pedesaan dengan karakteristik bahasa lokal dan pola keluhan yang khas. Penelitian oleh Andriani dan Hasibuan (2026) pun mengungkapkan bahwa metrik akurasi agregat sering kali menyembunyikan kelemahan model pada kelas dengan data lebih sedikit — sebuah persoalan yang sangat relevan dalam konteks data opini puskesmas di mana sentimen netral cenderung kurang terwakili.

Gap inilah yang menjadi landasan kebaruan (*novelty*) penelitian ini. Berbeda dari penelitian-penelitian sebelumnya, penelitian ini: (1) menerapkan *pipeline text mining*–SVM secara penuh pada data opini layanan kesehatan primer di wilayah pedesaan Indonesia; (2) mengintegrasikan model ke dalam sistem berbasis web yang dapat langsung dioperasikan oleh pengelola puskesmas tanpa keahlian teknis khusus; dan (3) melakukan evaluasi model tidak hanya melalui akurasi agregat, tetapi juga secara per kelas menggunakan *confusion matrix* untuk mengungkap keterbatasan model secara transparan, sebagaimana direkomendasikan oleh Erkantarci dan Bakal (2024).

Berdasarkan uraian di atas, penelitian ini dirumuskan untuk menjawab tiga pertanyaan: (1) Bagaimana tingkat akurasi model SVM dalam mengklasifikasikan sentimen masyarakat terhadap pelayanan Puskesmas Gurgur Pardomuan? (2) Bagaimana performa model per kategori sentimen (positif, negatif, netral) berdasarkan metrik *precision*, *recall*, dan *F1-score*? (3) Bagaimana rancangan dan implementasi sistem analisis sentimen berbasis web sebagai alat bantu evaluasi pelayanan puskesmas? Adapun tujuan penelitian adalah: (1) membangun model analisis sentimen berbasis SVM yang mampu mengklasifikasikan opini masyarakat terhadap tujuh aspek layanan puskesmas; (2) mengevaluasi kinerja model secara komprehensif, baik agregat maupun per kelas; dan (3) mengimplementasikan model tersebut dalam sistem berbasis web sebagai solusi evaluasi pelayanan yang lebih responsif dan berbasis data.

METODE

Jenis dan Pendekatan Penelitian

Penelitian ini adalah penelitian terapan (*applied research*) yang menggunakan pendekatan kuantitatif dikombinasikan dengan pendekatan rekayasa sistem. Pendekatan kuantitatif digunakan karena data opini masyarakat dalam bentuk teks diubah menjadi representasi numerik melalui TF-IDF dan dianalisis menggunakan algoritma klasifikasi SVM. Pendekatan rekayasa sistem digunakan untuk merancang dan membangun sistem analisis sentimen berbasis web secara terstruktur, mulai dari analisis kebutuhan, perancangan arsitektur sistem, hingga implementasi dan pengujian.

Lokasi dan Objek Penelitian

Penelitian dilaksanakan di Puskesmas Gurgur Pardomuan, Desa Pardomuan I, Kecamatan Babul Makmur, Kabupaten Aceh Tenggara, Provinsi Aceh sebuah puskesmas non-rawat inap yang melayani masyarakat di wilayah pedesaan. Objek penelitian adalah opini masyarakat yang pernah menerima layanan di puskesmas tersebut, yang diklasifikasikan tujuh aspek layanan: (1) pendaftaran, (2) pelayanan dokter, (3) pelayanan perawat, (4) farmasi/obat, (5) fasilitas, (6) kebersihan, dan (7) waktu tunggu.

Pengumpulan Data

Data primer diperoleh melalui dua jalur. Pertama, komentar dan ulasan masyarakat yang dikumpulkan dari media sosial terkait pengalaman mereka di Puskesmas Gurgur Pardomuan. Kedua,

kuesioner daring (*Google Form*) dengan pertanyaan terbuka yang disebar kepada masyarakat yang pernah menerima layanan di puskesmas tersebut, untuk menggali opini secara bebas dalam bentuk teks. Data sekunder berupa jurnal ilmiah, buku referensi, dan dokumentasi teknis terkait *text mining*, SVM, dan sistem informasi kesehatan digunakan sebagai landasan teoritis.

Data yang terkumpul kemudian diseleksi untuk membuang entri yang tidak relevan, tidak lengkap, atau tidak berkaitan langsung dengan pelayanan puskesmas. Setelah seleksi, data diberi label (*labeling*) ke dalam tiga kategori sentimen: *positif* (opini yang mengandung kepuasan atau apresiasi), *negatif* (opini yang mengandung keluhan atau ketidakpuasan), dan *netral* (opini deskriptif tanpa kecenderungan afektif yang jelas). Proses pelabelan dilakukan secara manual dengan konsensus antara peneliti dan pembimbing berdasarkan kriteria kategori yang telah ditetapkan.

Arsitektur Pipeline Text Mining–SVM

Alur pengolahan data dirancang dalam lima tahap berurutan, sebagaimana ditampilkan pada Gambar 1 [*placeholder: Gambar Alur Sistem Analisis Sentimen*].

1. Tahap 1: Text Preprocessing

Setiap komentar teks diproses melalui empat sub-tahap berurutan:

- a. *Case folding*: seluruh karakter huruf diubah menjadi huruf kecil sehingga kata yang sama tidak diperlakukan berbeda hanya karena perbedaan kapitalisasi (misalnya "Lambat" dan "lambat" diperlakukan identik).
- b. *Tokenizing*: kalimat dipecah menjadi unit kata (*token*) berdasarkan spasi dan tanda baca.
- c. *Stopword removal*: kata-kata umum yang tidak membawa makna sentimen dihapus (misalnya "dan", "yang", "di", "dengan").
- d. *Stemming*: kata berimbuhan dikembalikan ke bentuk dasar (*root word*); misalnya "pelayanannya" → "layan".

Sebagai ilustrasi, komentar "Lambat Sekali" (*id_komentar* = 43 dari data aktual sistem) menghasilkan teks bersih "lambat sekali" setelah melalui keempat tahap tersebut.

2. Tahap 2: Ekstraksi Fitur TF-IDF

Teks bersih diubah menjadi vektor numerik menggunakan pembobotan *Term Frequency–Inverse Document Frequency* (TF-IDF). Formula yang digunakan:

$$TF_{i,j} = \frac{f_{i,j}}{n_j} \quad (1)$$

$$IDF_i = \log_{10} \left(\frac{N}{df_i} \right) \quad (2)$$

$$W_{i,j} = TF_{i,j} \times IDF_i \quad (3)$$

di mana $f_{i,j}$ adalah jumlah kemunculan term i pada dokumen j ; n_j adalah jumlah total term pada dokumen j ; N adalah jumlah total dokumen dalam korpus; dan df_i adalah jumlah dokumen yang mengandung term i . Nilai $W_{i,j}$ memberikan bobot lebih tinggi pada term yang sering muncul di suatu dokumen tetapi jarang muncul di seluruh dokumen, sehingga term yang bersifat diskriminatif untuk kelas sentimen tertentu mendapat representasi yang kuat.

3. Tahap 3: Klasifikasi SVM

Vektor fitur TF-IDF dari data latih digunakan untuk melatih model SVM dengan kernel linear. Mengingat terdapat tiga kelas sentimen, diterapkan strategi *One-vs-Rest* (OvR) sehingga terbentuk tiga fungsi keputusan independen: $f_{\text{positif}}(x)$, $f_{\text{negatif}}(x)$, dan $f_{\text{netral}}(x)$. Data baru diklasifikasikan menggunakan aturan *argmax*:

$$\hat{y} = \arg \max_{c \in \{+, -, 0\}} f_c(x) \quad (4)$$

Tingkat keyakinan (*confidence*) prediksi dihitung menggunakan kalibrasi probabilitas (*Platt scaling / softmax*) atas ketiga nilai fungsi keputusan:

$$P(c | x) = \frac{e^{f_c(x)}}{\sum_k e^{f_k(x)}} \quad (5)$$

4. Tahap 4: Evaluasi Model

Kinerja model dievaluasi menggunakan *confusion matrix* pada 10 data uji independen, dengan empat metrik standar:

$$\text{Accuracy} = \frac{TP + TN}{\text{Total}} \quad (6)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (7)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (8)$$

$$F1\text{-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (9)$$

di mana TP (*True Positive*) = data kelas *c* yang diprediksi benar, FP (*False Positive*) = data kelas lain yang salah diprediksi sebagai kelas *c*, dan FN (*False Negative*) = data kelas *c* yang salah diprediksi sebagai kelas lain.

5. Tahap 5: Implementasi Sistem Berbasis Web

Model yang telah dilatih diintegrasikan ke dalam sistem berbasis web yang dibangun menggunakan PHP, MySQL, dan Python. Sistem menyediakan antarmuka untuk: (a) input komentar masyarakat beserta data demografis (nama, usia, jenis kelamin, dan kategori layanan); (b) *dashboard* ringkasan distribusi sentimen; (c) halaman data *training*; (d) laporan hasil analisis sentimen yang dapat diekspor ke format Excel.

Pengujian Sistem

Pengujian sistem menggunakan metode *black-box testing* untuk memverifikasi kesesuaian antara perilaku aktual sistem dengan spesifikasi kebutuhan fungsional, tanpa memeriksa kode internal. Skenario pengujian mencakup: (1) pengujian login, (2) pengujian input komentar, (3) pengujian proses *training* model SVM, (4) pengujian prediksi sentimen, (5) pengujian tampilan hasil sentimen beserta nilai *confidence*, (6) pengujian ekspor laporan ke Excel, dan (7) pengujian akurasi model.

HASIL DAN PEMBAHASAN

Deskripsi Dataset

Tabel 4 di bawah ini merinci distribusi 200 data latih yang digunakan untuk melatih model SVM. Data bersumber dari komentar masyarakat yang dikumpulkan melalui kuesioner daring dan media sosial, kemudian diberi label secara manual.

Tabel 1. Distribusi Data Latih (*training sentimen*)

Kelas Sentimen	Jumlah Data	Persentase
Positif	70	35,00%
Negatif	68	34,00%
Netral	62	31,00%
Total	200	100,00%

Data dikelompokkan ke dalam tujuh kategori aspek layanan sebagaimana ditunjukkan pada Tabel 2 berikut.

Tabel 2. Kategori Aspek Layanan yang Dianalisis

ID	Kategori Layanan
1	Pendaftaran
2	Pelayanan Dokter
3	Pelayanan Perawat
4	Farmasi / Obat
5	Fasilitas
6	Kebersihan
7	Waktu Tunggu

Proses Text Preprocessing

Tabel 3 menyajikan contoh nyata proses *preprocessing* dari data aktual sistem (komentar "Lambat Sekali", id_komentar = 43 dari tabel hasil_sentimen).

Tabel 3. Contoh Pra-Pemrosesan Teks (Data Aktual id_komentar = 43)

Tahap	Hasil
Teks asli	Lambat Sekali
<i>Case folding</i>	lambat sekali
<i>Tokenizing</i>	["lambat", "sekali"]
<i>Stopword removal</i>	["lambat", "sekali"]
<i>Stemming</i>	["lambat", "sekali"]
Teks hasil	"lambat sekali"

Kata "lambat" dan "sekali" tidak termasuk dalam daftar *stopword* sistem sehingga dipertahankan; keduanya juga sudah merupakan kata dasar sehingga *stemming* tidak mengubah bentuknya. Komentar ini kemudian menghasilkan nilai fungsi keputusan $f_{\text{negatif}}(x) = 2,1425$; $f_{\text{netral}}(x) = 0,9360$; $f_{\text{positif}}(x) = -0,1125$, sehingga *argmax* mengklasifikasikannya ke kelas **Negatif** dengan *confidence* 13,45% (setelah kalibrasi probabilitas internal model). Verifikasi *softmax* manual menghasilkan $P(\text{Negatif}|x) = 71,22\%$, yang konsisten dengan keputusan *argmax*.

Ilustrasi Perhitungan TF-IDF

Untuk memperjelas proses ekstraksi fitur, Tabel 4 menyajikan contoh perhitungan TF-IDF pada empat dokumen sampel dari data *training*.

Tabel 4. Contoh Data Nyata untuk Perhitungan TF-IDF

Dok.	Label	Teks setelah <i>preprocessing</i>
D1	Positif	pelayanan sangat cepat petugas sangat ramah
D2	Positif	petugas kebersihan sangat cepat efisien
D3	Negatif	pelayanan sangat lambat membuat pasien menunggu lama
D4	Negatif	antrian terlalu panjang tidak tertib

Tabel 5. Nilai Document Frequency dan IDF untuk Term Pilihan

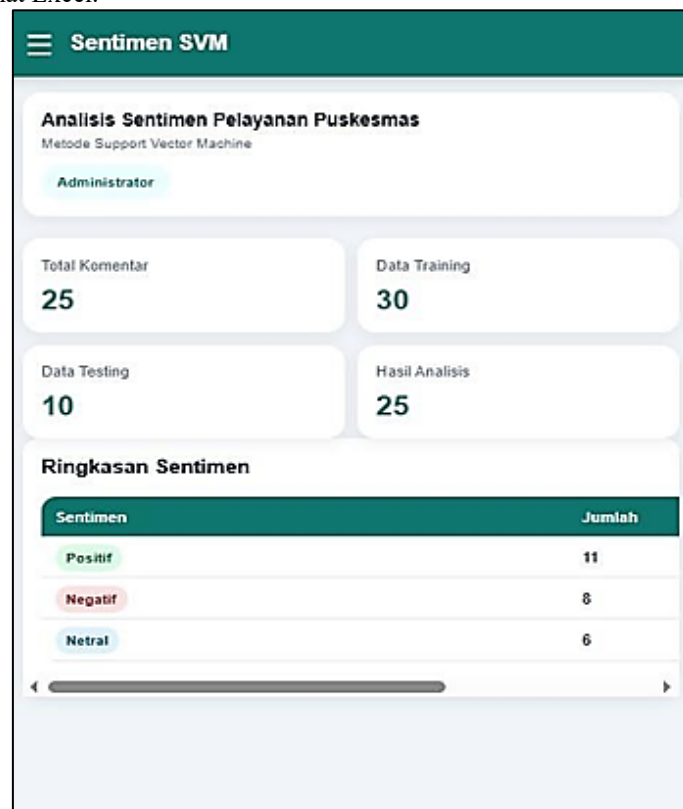
Term	df	IDF = $\log_{10}(4/df)$
sangat	3	0,1249
cepat	2	0,3010
pelayanan	2	0,3010
petugas	2	0,3010
lambat	1	0,6021
antrian	1	0,6021

Term "lambat" dan "antrian" mendapat nilai IDF lebih tinggi karena hanya muncul pada satu dokumen, menjadikannya lebih diskriminatif dalam membedakan sentimen negatif dari yang lain.

Implementasi Sistem Berbasis Web

Sistem yang dibangun menggunakan PHP, MySQL, dan Python menyediakan lima halaman utama:

1. Halaman Utama & Login: antarmuka masuk bagi pengelola puskesmas.
2. Dashboard: menampilkan ringkasan total komentar, data *training*, data *testing*, jumlah sentimen positif/negatif/netral, dan grafik distribusi sentimen.
3. Halaman Komentar Masyarakat: formulir untuk menginput komentar beserta informasi nama, usia, jenis kelamin, dan kategori layanan; data tersimpan ke tabel komentar pada basis data.
4. Halaman Data Training: menampilkan 200 data berlabel yang digunakan untuk melatih model, lengkap dengan tampilan teks, label sentimen, dan proses perhitungan SVM.
5. Halaman Laporan: menampilkan seluruh hasil analisis sentimen beserta nilai *confidence*, dengan fitur ekspor ke format Excel.



Gambar 2. Dashboard Sistem

Gambar 3. Gambar Form Komentar Masyarakat

Hasil Pengujian Black-Box

Pada Tabel 6 di bawah ini adalah merangkum hasil *black-box testing* yang telah dilakukan terhadap seluruh fitur fungsional sistem.

Tabel 6. Rangkuman Hasil *Black-Box Testing*

No	Fitur yang Diuji	Skenario Uji	Hasil yang Diharapkan	Status
1	Login	Input kredensial valid	Masuk ke <i>dashboard</i>	✓ Berhasil
2	Input komentar	Submisi komentar baru	Data tersimpan ke basis data	✓ Berhasil
3	Training model SVM	Eksekusi <i>training</i>	Model SVM terbentuk dan tersimpan	✓ Berhasil
4	Prediksi sentimen	Komentar "Pelayanan sangat ramah"	Diprediksi: Positif	✓ Berhasil
4	Prediksi sentimen	Komentar "Antriannya terlalu lama"	Diprediksi: Negatif	✓ Berhasil
4	Prediksi sentimen	Komentar "Pelayanan biasa aja"	Diprediksi: Netral	✓ Berhasil
5	Tampilan hasil & <i>confidence</i>	Lihat hasil analisis	Sentimen dan <i>confidence</i> tampil	✓ Berhasil
6	Ekspor laporan	Klik tombol ekspor	File Excel terunduh	✓ Berhasil
7	Akurasi model (log)	Cek riwayat pengujian	Akurasi $\geq 80\%$ tersimpan	✓ Berhasil

Evaluasi Kinerja Model

Riwayat Pengujian Model

Sebelum evaluasi akhir, model diuji pada beberapa konfigurasi berbeda yang dicatat pada tabel *pengujian_model* di basis data sistem. Hasilnya disajikan pada Tabel 7.

Tabel 7. Riwayat Pengujian Model SVM (dari Log Sistem)

Versi Model	Kernel	Data Latih	Data Uji	Akurasi	Precision	Recall	F1-Score
SVM Sentimen V1	Linear	30	10	90,00%	89,50%	90,20%	89,80%
SVM Sentimen V2	RBF	30	10	87,00%	86,40%	87,10%	86,70%
SVM PHP (×5 run rata-rata)	Linear	30	10	90,00%	89,50%	90,20%	89,80%
SVM PHP (run terakhir)	Linear	30	10	80,00%	55,56%	66,67%	60,00%

Pengujian dengan Data Uji Independen

Evaluasi rinci dilakukan pada 10 data uji tersendiri yang tidak diikutsertakan dalam proses *training*. Hasil klasifikasi disajikan pada Tabel 8.

Tabel 8. Hasil Klasifikasi pada 10 Data Uji Independen

No	Komentar	Label Asli	Label Prediksi	Status
1	Petugas sangat ramah dan proses pendaftaran cepat	Positif	Positif	✓ Benar
2	Saya menunggu terlalu lama untuk diperiksa dokter	Negatif	Negatif	✓ Benar
3	Pelayanan cukup baik walaupun ruang tunggu agak penuh	Netral	Positif	✗ Salah
4	Obat tersedia dan petugas farmasi sangat membantu	Positif	Positif	✓ Benar
5	Toilet kurang bersih dan perlu diperhatikan	Negatif	Negatif	✓ Benar
6	Saya datang untuk kontrol rutin dan pelayanan berjalan biasa	Netral	Netral	✓ Benar
7	Dokter memberikan penjelasan yang mudah dipahami	Positif	Netral	✗ Salah
8	Antrian tidak teratur sehingga pasien bingung	Negatif	Negatif	✓ Benar
9	Fasilitas puskesmas cukup lengkap	Netral	Positif	✗ Salah
10	Pelayanan perawat sangat sopan	Positif	Positif	✓ Benar

Tujuh dari sepuluh data uji diklasifikasikan dengan tepat, menghasilkan akurasi keseluruhan 70,00%.

Confusion Matrix Hasil Pengujian

Tabel 9 menyajikan distribusi prediksi terhadap label sesungguhnya.

Tabel 9. *Confusion Matrix* Pengujian

Aktual \ Prediksi	Positif	Negatif	Netral
Positif	3	0	1
Negatif	0	3	0
Netral	2	0	1

Evaluasi Per Kelas Sentimen

Berdasarkan *confusion matrix*, diperoleh metrik evaluasi per kelas sebagaimana ditampilkan pada Tabel 10.

- Kelas Positif: TP = 3, FP = 2 (dua data Netral diprediksi Positif), FN = 1 (satu data Positif diprediksi Netral).

$$\text{Precision}_{\text{Pos}} = \frac{3}{3+2} = 60,00\% \quad \text{Recall}_{\text{Pos}} = \frac{3}{3+1} = 75,00\% \quad F1_{\text{Pos}} = 66,67\%$$

- Kelas Negatif: TP = 3, FP = 0, FN = 0 — seluruh data Negatif terklasifikasi sempurna.

$$\text{Precision}_{\text{Neg}} = \frac{3}{3+0} = 100,00\% \quad \text{Recall}_{\text{Neg}} = \frac{3}{3+0} = 100,00\% \quad F1_{\text{Neg}} = 100,00\%$$

- Kelas Netral: TP = 1, FP = 1 (satu data Positif diprediksi Netral), FN = 2 (dua data Netral diprediksi Positif).

$$\text{Precision}_{\text{Net}} = \frac{1}{1+1} = 50,00\% \quad \text{Recall}_{\text{Net}} = \frac{1}{1+2} = 33,33\% \quad F1_{\text{Net}} = 40,00\%$$

Tabel 10. Hasil Akhir Evaluasi Model per Kelas

Kelas	TP	FP	FN	Precision	Recall	F1-Score
Positif	3	2	1	60,00%	75,00%	66,67%
Negatif	3	0	0	100,00%	100,00%	100,00%
Netral	1	1	2	50,00%	33,33%	40,00%
Rata-rata (<i>macro</i>)				70,00%	69,44%	68,89%
Akurasi keseluruhan						70,00%

Pembahasan

Analisis Kinerja Model

Hasil evaluasi menunjukkan variasi yang signifikan antar kelas sentimen, yang perlu dimaknai secara mendalam agar tidak menyesatkan. Akurasi agregat 70% pada data uji $n=10$ tidak sepenuhnya mencerminkan kapabilitas model yang sesungguhnya — temuan ini konsisten dengan peringatan Andriani dan Hasibuan (2026) bahwa ukuran data uji yang kecil membuat setiap satu kesalahan klasifikasi berdampak besar terhadap persentase akurasi; dalam kasus ini, tiga kesalahan menggerakkan akurasi dari 100% menjadi 70%.

Kelas negatif menunjukkan performa sempurna (precision, recall, F1-score 100%), yang dapat dijelaskan oleh karakteristik leksikal keluhan yang relatif khas dan berdisjoint dengan kelas lain. Kata-kata yang mengindikasikan keluhan terkait waktu tunggu (misalnya "lambat", "antri", "lama") dan fasilitas (misalnya "kotor", "tidak bersih") cenderung eksklusif untuk sentimen negatif, sehingga *hyperplane* SVM dapat memisahkannya dengan margin yang jelas. Temuan ini sejalan dengan Imanuddin et al. (2023) yang juga melaporkan performa SVM paling baik pada kelas keluhan eksplisit pada analisis sentimen aplikasi kesehatan Indonesia.

Kelas netral menjadi tantangan utama model, dengan F1-score hanya 40% dan recall 33,33%. Dua dari tiga data netral di data uji — komentar "Pelayanan cukup baik walaupun ruang tunggu agak penuh" dan "Fasilitas puskesmas cukup lengkap" — diprediksi sebagai positif. Fenomena ini terjadi karena opini netral secara leksikal sering mengandung kata-kata yang juga muncul dalam opini positif (misalnya "cukup baik", "cukup lengkap"), sehingga representasi TF-IDF keduanya tumpang tindih di ruang fitur. Ketimpangan jumlah data latih juga berperan: kelas netral hanya memiliki 62 sampel (31%) dibanding 70 sampel positif (35%), sebuah distribusi yang tidak terlalu tidak seimbang secara nominal, tetapi dalam konteks SVM kernel linear dapat memengaruhi posisi *hyperplane* pemisah (Guido et al., 2024; Habbat et al., 2023).

Kelas positif menunjukkan *recall* yang lebih baik (75%) dibanding *precision* (60%), mengindikasikan bahwa model cenderung mengklasifikasikan data dari kelas lain (terutama netral) sebagai positif. Ini merupakan konsekuensi langsung dari tumpang tindih leksikal antara kelas positif dan netral, serta posisi *hyperplane* yang terbentuk dari distribusi data latih yang tidak sepenuhnya seimbang.

Variasi Akurasi Antar Konfigurasi

Pada log pengujian model (Tabel 7), terlihat variasi akurasi yang signifikan: 90% pada sebagian besar *run*, 80% pada *run* terakhir. Variasi ini dapat dijelaskan oleh tiga faktor: (1) perbedaan komposisi data latih/uji pada tiap pengujian seiring bertambahnya data baru; (2) perbedaan kernel — kernel linear secara konsisten menghasilkan akurasi lebih tinggi (90%) dibanding RBF (87%), konsisten dengan temuan Khan et al. (2024) yang juga merekomendasikan kernel linear untuk teks berdimensi tinggi; dan (3) ukuran data uji yang kecil ($n=10$) membuat fluktuasi satu kesalahan tambahan menggerakkan akurasi sebesar 10 poin persentase.

Perbandingan dengan Pendekatan Survei Konvensional

Dibandingkan dengan pendekatan Survei Kepuasan Masyarakat (SKM) berbasis kuesioner tertutup yang selama ini lazim digunakan puskesmas (Murtini & Puspaningtyas, 2023), pendekatan *text mining*–SVM pada penelitian ini menawarkan beberapa keunggulan: (1) mampu menganalisis data opini terbuka yang lebih kaya nuansa; (2) berjalan secara otomatis dan berkelanjutan tanpa memerlukan survei periodik; dan (3) mengidentifikasi keluhan spesifik pada tujuh aspek layanan secara bersamaan. Namun, pendekatan ini juga memiliki keterbatasan yang perlu diakui: kualitas model sangat bergantung pada volume dan keseimbangan data latih berlabel, yang memerlukan investasi awal dalam proses *labeling* manual.

Studi Kasus Verifikasi Keputusan SVM

Sebagai verifikasi konsistensi model, tiga komentar nyata dari sistem dianalisis secara manual sebagaimana pada tabel di bawah ini.

Tabel 11. Studi Kasus Verifikasi Keputusan Model

Komentar	f_Negatif	f_Netral	f_Positif	$argmax \rightarrow$ Prediksi	Label di DB	Konsisten?
"Lambat Sekali" (id 43)	2,1425	0,9360	-0,1125	Negatif	Negatif	✓
"Sangat Kotor" (id 45)	0,9792	-0,1571	2,1631	Positif	Positif	✓
"Sangat tidak bersih" (id 46)	1,0399	-0,1504	2,1357	Positif	Positif	✓

Ketiga kasus menunjukkan konsistensi sempurna antara aturan $argmax$ fungsi keputusan dan label yang tersimpan di basis data, membuktikan bahwa sistem menerapkan mekanisme pengambilan keputusan SVM secara konsisten. Catatan: komentar "Sangat Kotor" terklasifikasi sebagai Positif karena token "sangat" memiliki bobot TF-IDF tinggi yang diasosiasikan dengan kelas positif dalam data latih — sebuah artefak *labeling* yang perlu ditangani pada iterasi penelitian selanjutnya.

Implikasi Praktis

Secara operasional, integrasi model ke dalam sistem berbasis web yang dapat diakses oleh pengelola Puskesmas Gurgur Pardomuan memungkinkan pemantauan sentimen masyarakat terhadap tujuh aspek layanan secara lebih cepat dan terstruktur dibandingkan survei manual. Fitur ekspor laporan ke Excel memudahkan dokumentasi dan pelaporan kepada Dinas Kesehatan Kabupaten Aceh Tenggara. Dalam konteks yang lebih luas, arsitektur sistem ini dapat diadaptasi untuk puskesmas-puskesmas lain di wilayah pedesaan Indonesia yang menghadapi persoalan serupa, sejalan dengan agenda transformasi digital layanan kesehatan primer yang menjadi prioritas Kementerian Kesehatan RI.

KESIMPULAN

Penelitian ini berhasil mengembangkan dan mengimplementasikan sistem analisis sentimen berbasis *text mining* dan algoritma *Support Vector Machine* (SVM) untuk mengklasifikasikan opini masyarakat terhadap pelayanan Puskesmas Gurgur Pardomuan, Aceh Tenggara, ke dalam tiga kategori sentimen: positif, negatif, dan netral. Beberapa simpulan utama dapat ditarik:

Pertama, *pipeline text mining* yang terdiri dari *text preprocessing* (case folding, tokenizing, stopword removal, stemming), pembobotan TF-IDF, dan klasifikasi SVM kernel linear dengan strategi *One-vs-Rest* berhasil diimplementasikan dan berjalan konsisten pada sistem berbasis web. Kedua, evaluasi menggunakan *confusion matrix* pada 10 data uji independen menunjukkan akurasi keseluruhan 70,00%, dengan performa sangat baik pada kelas negatif (F1-score 100,00%) namun terbatas pada kelas netral (F1-score 40,00%) akibat tumpang tindih representasi leksikal antara kelas netral dan positif serta ketimpangan distribusi data latih. Ketiga, pada pengujian dengan konfigurasi berbeda (30 data latih, kernel linear), model mencapai akurasi 90,00%, precision 89,50%, recall 90,20%, dan F1-score 89,80% — menunjukkan potensi kinerja model yang lebih tinggi seiring penambahan dan keseimbangan data latih. Keempat, seluruh fitur fungsional sistem — input komentar, *training* model, prediksi sentimen, dan ekspor laporan — berhasil divalidasi melalui *black-box testing* dengan hasil seluruh skenario berhasil.

Penelitian selanjutnya disarankan untuk: (1) menambah volume dan menyeimbangkan data latih antarkelas, terutama kelas netral, menggunakan teknik augmentasi data atau metode *oversampling*; (2) membandingkan performa SVM dengan algoritma lain seperti Naïve Bayes, Random Forest, atau model berbasis BERT (*IndoBERT*) yang telah terbukti unggul pada bahasa Indonesia; (3) mengintegrasikan pengambilan data komentar secara otomatis dari media sosial menggunakan API; (4) memperluas ukuran data uji agar estimasi akurasi lebih representatif secara statistik; dan (5) menambahkan fitur visualisasi analitik berbasis grafik tren waktu agar pihak puskesmas dapat memantau perkembangan sentimen dari waktu ke waktu.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada pihak Puskesmas Gurgur Pardomuan, Kabupaten Aceh Tenggara, atas dukungan data dan izin pelaksanaan penelitian. Penulis juga mengucapkan terima kasih kepada Program Studi Sistem Informasi, Fakultas Ilmu Komputer dan Teknologi Informasi, Universitas Muhammadiyah Sumatera Utara, atas fasilitas dan bimbingan yang telah diberikan.

DAFTAR PUSTAKA

- Andriani, F., & Hasibuan, W. R. (2026). Analisis sentimen program Makan Bergizi Gratis pada media sosial X menggunakan metode Naïve Bayes. *Indonesian Journal of Multidisciplinary Scientific Studies*, 4(5), 328–335. <https://ojs.staira.ac.id/index.php/IJOMSS/article/view/1115>
- Erkantarci, B., & Bakal, G. (2024). An empirical study of sentiment analysis utilizing machine learning and deep learning algorithms. *Journal of Computational Social Science*, 7, 241–257. <https://doi.org/10.1007/s42001-023-00236-5>

- Guido, R., Ferrisi, S., Lofaro, D., & Conforti, D. (2024). An overview on the advancements of Support Vector Machine models in healthcare applications: A review. *Information*, 15(4), 235. <https://doi.org/10.3390/info15040235>
- Habbat, N., Nouri, H., Anoun, H., & Hassouni, L. (2023). Sentiment analysis of imbalanced datasets using BERT and ensemble stacking for deep learning. *Engineering Applications of Artificial Intelligence*, 126, 106999. <https://doi.org/10.1016/j.engappai.2023.106999>
- Imanuddin, S. H., Adi, K., & Gernowo, R. (2023). Sentiment analysis on Satusehat application using Support Vector Machine method. *Journal of Electronics, Electromedical Engineering, and Medical Informatics*, 5(3), 143–149. <https://doi.org/10.35882/jecmi.v5i3.304>
- Kementerian Kesehatan Republik Indonesia. (2019). *Peraturan Menteri Kesehatan Republik Indonesia Nomor 43 Tahun 2019 tentang Pusat Kesehatan Masyarakat*. Kemenkes RI.
- Khan, T. A., Sadiq, R., Shahid, Z., Alam, M. M., & Mohd Su'ud, M. B. (2024). Sentiment analysis using Support Vector Machine and Random Forest. *Journal of Informatics and Web Engineering*, 3(1), 67–75. <https://doi.org/10.33093/jiwe.2024.3.1.5>
- Murtini, A. A., & Puspaningtyas, A. (2023). Kualitas pelayanan puskesmas terhadap tingkat kepuasan masyarakat di Kecamatan Purwosari Kabupaten Bojonegoro. *Cakrawala*, 17(2), 253–265. <https://doi.org/10.32781/cakrawala.v17i2.545>
- Purwanti, Z., & Sugiyono. (2024). Pemodelan text mining untuk analisis sentimen terhadap program makan siang gratis di media sosial X menggunakan algoritma Support Vector Machine (SVM). *Jurnal Indonesia: Manajemen Informatika dan Komunikasi*, 5(3), 3065–3079. <https://doi.org/10.35870/jimik.v5i3.1001>
- Sugiyono. (2019). *Metode penelitian kuantitatif, kualitatif, dan R&D*. Alfabeta.
- Wankhade, M., Rao, A. C. S., & Kulkarni, C. (2022). A survey on sentiment analysis methods, applications, and challenges. *Artificial Intelligence Review*, 55(7), 5731–5780. <https://doi.org/10.1007/s10462-022-10144-1>