



Analisis Sentimen Ulasan *Google Maps* Berbasis *Natural Language Processing* (NLP) Sebagai Gambaran Kualitas Pelayanan RS Mata Makassar

Idarwati Mustafa, Muh. Ashhar Bustan, Nur Assyurthy, Munir

idarwati.mustafa@gmail.com

The work is licensed under a Creative Commons Attribution License (CC BY-SA 4.0)

How to Cite:

Mustafa, I., Bustan, M. A., Assyurthy, N., & Munir. Analisis Sentimen Ulasan Google Maps Berbasis Natural Language Processing (NLP) Sebagai Gambaran Kualitas Pelayanan RS Mata Makassar. Jurnal of Ophthalmology 2026, volume 3(1), halaman (38–44).

<https://doi.org/10.63670/mata.v2i1.50>

ABSTRACT

Introduction: Google Maps reviews provide spontaneous and continuous patient feedback, but their potential for assessing specific dimensions of hospital service quality remains underused. This study aimed to evaluate service quality at RS Mata Makassar through Natural Language Processing (NLP)-based sentiment analysis and the SERVQUAL framework. **Methods:** This quantitative descriptive study analyzed all eligible Google Maps reviews published from January 2024 to March 2026. Data were collected through manual scraping and processed using cleaning, case folding, normalization, and stopword removal. Sentiment was classified using a rule-based approach into positive, negative, or neutral categories, while service aspects were mapped multilabel to Tangibles, Reliability, Responsiveness, Assurance, and Empathy. **Results:** The analysis generated 1,495 aspect–sentiment assignments: 1,416 (94.72%) positive, 44 (2.94%) neutral, and 35 (2.34%) negative. Reliability represented the most frequently discussed dimension and was predominantly associated with positive experiences, whereas Empathy demonstrated the highest proportion of positive sentiment. Nevertheless, negative feedback was disproportionately concentrated in the Responsiveness dimension, mainly concerning waiting time, queues, delays, proactive service communication as priority areas for improvement. **Conclusion:** Google Maps reviews have the potential to complement patient satisfaction evaluations and support continuous monitoring of hospital service quality. Reviews of RS Mata Makassar were predominantly positive, particularly in the Reliability and Empathy dimensions. However, Responsiveness emerged as the main priority for improvement because it had the highest proportion of negative sentiment and accounted for nearly half of all negative aspect–sentiment assignments.

Keywords: Google Maps; Hospital Service Quality; Natural Language Processing; Sentiment Analysis; SERVQUAL.

PENDAHULUAN

Pengalaman pasien merupakan salah satu sumber informasi penting dalam evaluasi mutu pelayanan karena menggambarkan bagaimana pasien menilai akses, proses, interaksi, dan lingkungan pelayanan. Survei kepuasan terstruktur tetap dibutuhkan, tetapi pertanyaan yang telah ditetapkan sebelumnya dapat membatasi jenis persoalan yang dapat disampaikan pasien. Sebaliknya, ulasan daring memungkinkan pasien menuliskan pengalaman secara spontan dan kontekstual. Umpan balik daring relatif mudah diakses dan berbiaya rendah, tetapi pemberi ulasan tidak selalu mewakili seluruh populasi pasien; karena itu, data ini lebih tepat digunakan sebagai pelengkap survei dan data operasional [1].



Besarnya volume dan bentuk teks yang tidak terstruktur membuat ulasan daring sulit dianalisis secara manual dan konsisten. *Natural Language Processing* (NLP) memungkinkan analisis sentimen, pemodelan topik, dan klasifikasi teks untuk mengekstraksi pola, tema, serta area perbaikan dari umpan balik pasien [2]. Dalam kerangka tersebut, *Aspect-Based Sentiment Analysis* (ABSA) mengaitkan sentimen dengan aspek yang dibicarakan sehingga satu ulasan dapat memuat beberapa tema pelayanan. Analisis ulasan rumah sakit juga menunjukkan bahwa pengolahan teks otomatis dapat menghasilkan tema yang dapat ditindaklanjuti, tetapi temuannya lebih tepat digunakan untuk membentuk hipotesis bagi investigasi lanjutan [3].

Untuk menerjemahkan sentimen menjadi informasi mutu pelayanan, penelitian ini menggunakan kerangka SERVQUAL yang mencakup *Tangibles*, *Reliability*, *Responsiveness*, *Assurance*, dan *Empathy* [4]. Pemetaan ke lima dimensi tersebut diperlukan karena *rating* bintang atau sentimen keseluruhan dapat menyembunyikan perbedaan antara aspek yang paling diapresiasi dan aspek yang memiliki konsentrasi keluhan tertinggi. Dengan demikian, volume penyebutan suatu dimensi perlu dibaca bersama proporsi sentimen negatifnya.

Dalam konteks RS Mata Makassar, bukti empiris mengenai bagaimana pengalaman spontan pasien terdistribusi pada lima dimensi SERVQUAL masih terbatas. Penilaian berbasis *rating* saja belum dapat menjelaskan aspek yang paling diapresiasi maupun keluhan yang paling memerlukan tindakan. Berangkat dari kesenjangan tersebut, penelitian ini bertujuan menganalisis sentimen ulasan Google Maps RS Mata Makassar menggunakan ABSA, memetakan setiap penetapan aspek sentimen ke dalam dimensi SERVQUAL, serta menentukan prioritas perbaikan dengan mempertimbangkan volume penyebutan, proporsi sentimen negatif, dan tema keluhan. Penelitian ini diharapkan menghasilkan mekanisme deteksi awal berbasis eWOM yang melengkapi survei kepuasan, data komplain, dan data operasional untuk mendukung perbaikan mutu yang lebih responsif dan terarah.

METODE

Penelitian ini menggunakan desain deskriptif kuantitatif dengan pendekatan *Natural Language Processing* (NLP). Desain tersebut digunakan untuk menggambarkan distribusi sentimen dan penetapan aspek pelayanan secara sistematis. Analisis disajikan dalam bentuk frekuensi, proporsi, serta perbandingan distribusi sentimen antar dimensi SERVQUAL.

ABSA digunakan dalam kerangka NLP untuk mengidentifikasi sentimen pada aspek pelayanan yang spesifik, bukan hanya menentukan sentimen umum suatu ulasan [2]. Setiap ulasan dipetakan ke dalam lima dimensi SERVQUAL, yaitu *Tangibles* (Bukti Fisik), *Reliability* (Keandalan), *Responsiveness* (Daya Tanggap), *Assurance* (Jaminan), dan *Empathy* (Empati). Satu ulasan dapat dipetakan ke lebih dari satu dimensi apabila memuat beberapa aspek pelayanan. Sumber data adalah ulasan pada laman Google Maps “Rumah Sakit Mata



Kemenkes RI Makassar” yang dipublikasikan pada Januari 2024–Maret 2026. Pengumpulan data dilakukan secara daring pada April–Mei 2026.

Populasi penelitian adalah seluruh ulasan yang dipublikasikan pada akun Google Maps RS Mata Makassar selama Januari 2024–Maret 2026. Sampel ditentukan dengan total sampling terhadap seluruh ulasan yang memenuhi kriteria. Perangkat lunak berbasis Python digunakan sebagai instrumen pengumpulan dan analisis. Data dikumpulkan melalui *scraping* manual dengan atribut berupa teks ulasan, waktu publikasi, dan rating bintang. Prosedur analisis terdiri atas empat tahap. Pertama, pengumpulan data (*data crawling*). Kedua, pra-pemrosesan teks yang mencakup cleaning untuk menghapus emoji, tanda baca berlebih, dan angka yang tidak relevan; *case folding* untuk mengubah huruf menjadi huruf kecil; normalisasi istilah; serta *stopword removal* untuk menghapus kata yang tidak memengaruhi analisis sentimen. Ketiga, pelabelan dimensi (*dimension tagging*). Sistem memindai kata kunci SERVQUAL dan mengategorikan teks ke dimensi kualitas pelayanan yang sesuai. Pelabelan dilakukan secara multilabel apabila satu ulasan memuat kata kunci dari beberapa dimensi. Keempat, analisis sentimen dengan teknik *scoring*. Setiap penetapan aspek diberikan skor berdasarkan kata atau frasa yang mengandung makna positif maupun negatif. Skor >0 dikategorikan positif, skor <0 dikategorikan negatif, dan skor $=0$ dikategorikan netral.

HASIL PENELITIAN

Analisis menghasilkan 1.495 penetapan aspek sentimen dari ulasan Google Maps RS Mata Makassar. Sebanyak 1.416 (94,72%) penetapan diklasifikasikan positif, 35 (2,34%) negatif, dan 44 (2,94%) netral. Jumlah tersebut bukan jumlah ulasan karena satu ulasan dapat memuat lebih dari satu aspek SERVQUAL.

Tabel 1 Distribusi Kategori Sentimen Berdasarkan Dimensi SERVQUAL

Dimensi SERVQUAL	Kategori Sentimen						Total Aspek
	Positif		Negatif		Netral		
	n	%	n	%	n	%	
Tangibles	189	92,65	6	2,94	9	4,41	204
Reliability	467	95,50	9	1,84	13	2,66	489
Responsiveness	122	83,56	16	10,96	8	5,48	146
Assurance	187	94,92	2	1,02	8	4,06	197
Empathy	451	98,26	2	0,44	6	1,31	459
Total	1.416	94,72	35	2,34	44	2,94	1.495

Dominasi penetapan aspek sentimen positif menunjukkan arah yang konsisten dengan distribusi rating yang didominasi bintang 5. Namun, kedua indikator tersebut tidak identik: *rating*



merupakan penilaian numerik pada tingkat ulasan, sedangkan analisis sentimen dapat menghasilkan lebih dari satu penetapan aspek sentimen dari satu ulasan.

Seluruh dimensi SERVQUAL didominasi sentimen positif. *Empathy* memiliki proporsi positif tertinggi, yaitu 98,26%. *Responsiveness* memiliki jumlah dan proporsi negatif tertinggi, yaitu 16 penetapan (10,96%) dari seluruh penetapan pada dimensi tersebut.

Tabel 2. Aspek Yang Teridentifikasi Berdasarkan Dimensi SERVQUAL Menggunakan *Natural Language Processing*

Dimensi	Aspek yang ditemukan	Kata atau frasa yang muncul
<i>Responsiveness</i>	Waktu tunggu, antrean dan keterlambatan pelayanan	Menunggu, lambat, lama, antre
<i>Reliability</i>	Administrasi, pelayanan BPJS dan proses rujukan	BPJS, rujukan, dipingpong
<i>Tangibles</i>	Detail fisik atau fasilitas selama pelayanan	Fasilitas, kondisi fisik, kuku tajam
<i>Assurance</i>	Interaksi dan perilaku petugas tertentu	Judes, kasar

Hasil identifikasi kata dan frasa pada ulasan negatif menunjukkan bahwa keluhan dalam dimensi *Responsiveness* terutama ditandai oleh kata “menunggu”, “lambat”, “lama”, dan “antre”. Pada dimensi *Reliability*, keluhan berkaitan dengan proses administrasi, pelayanan BPJS, sistem rujukan, serta proses pelayanan yang dinilai berulang atau berbelit. Keluhan pada dimensi *Tangibles* dan *Assurance* mencakup detail fisik saat pelayanan dan interaksi dengan petugas tertentu.

PEMBAHASAN

Penelitian ini menunjukkan dua temuan utama. Pertama, pengalaman yang dituliskan pasien secara keseluruhan sangat positif, terutama pada *Reliability* dan *Empathy*. Kedua, keluhan terkonsentrasi pada *Responsiveness*, meskipun dimensi ini bukan yang paling banyak disebut. Dengan menggunakan kerangka SERVQUAL [4], penelitian ini membaca volume penyebutan bersama proporsi negatif dan kontribusinya terhadap seluruh keluhan. Pendekatan tersebut mengembangkan pemeringkatan dimensi menjadi penyaringan prioritas perbaikan yang lebih operasional.

Dominasi sentimen positif dan rating bintang 5 menunjukkan konsistensi arah antara narasi dan penilaian numerik. Namun, angka ini tidak boleh ditafsirkan sebagai estimasi prevalensi kepuasan seluruh pasien karena pemberi ulasan daring merupakan kelompok yang memilih sendiri untuk menulis dan dapat berbeda dari keseluruhan populasi pasien. Kajian tentang umpan balik pasien daring menempatkan sumber ini sebagai mekanisme yang murah, mudah diakses, dan kaya konteks, tetapi tetap perlu digunakan sebagai pelengkap, bukan pengganti,



survei terstruktur dan data operasional [1]. Studi terbaru juga menunjukkan bahwa komentar daring dapat memuat informasi yang dapat ditindaklanjuti; ulasan negatif atau campuran cenderung lebih dapat ditindaklanjuti daripada ulasan positif [5].

Kekuatan pada *Reliability* dan *Empathy* tercermin dalam penyebutan pelayanan yang baik, petugas yang ramah dan membantu, serta lingkungan yang bersih dan nyaman. Dalam kerangka SERVQUAL, *Reliability* menggambarkan kemampuan memberikan pelayanan secara konsisten dan dapat dipercaya, sedangkan *Empathy* menggambarkan perhatian individual kepada pasien [4]. Karena data penelitian ini bersifat observasional dan tidak memuat pengukuran klinis, temuan tersebut perlu dipahami sebagai persepsi yang diungkapkan pasien, bukan bukti langsung mengenai efektivitas klinis. Meski demikian, kekuatan interpersonal ini merupakan modal organisasi yang dapat dipertahankan melalui umpan balik positif kepada unit, orientasi pegawai, penguatan komunikasi, dan promosi rumah sakit berbasis bukti.

Responsiveness menjadi titik risiko karena 10,96% penetapan dalam dimensi ini bersentimen negatif dan mencakup 45,71% dari seluruh penetapan negatif. Tema menunggu, antrean, dan keterlambatan sesuai dengan penelitian yang menunjukkan bahwa kepuasan pasien tidak hanya dipengaruhi waktu tunggu aktual, tetapi juga waktu yang dipersepsikan serta kesenjangan antara waktu yang diharapkan dan waktu aktual [6]. Pada pelayanan mata, siklus *Plan-Do-Study-Act* (PDSA) dan pencatatan waktu pada setiap titik layanan telah digunakan untuk mengukur waktu tunggu, waktu pelayanan, serta waktu menganggur dan mengidentifikasi peluang perbaikan alur pasien [7].

Keluhan mungkin tidak semata-mata berkaitan dengan lamanya menunggu, tetapi juga dengan menunggu tanpa penjelasan. Penelitian terdahulu menemukan bahwa kualitas komunikasi dapat memoderasi dan mengurangi dampak negatif waktu tunggu yang dipersepsikan terhadap kepuasan pasien [8]. Hal yang sama berlaku pada pemeriksaan berulang. Pengulangan dapat memiliki alasan klinis, keselamatan, atau validasi hasil, sehingga ulasan pasien tidak cukup untuk menyatakan bahwa pemeriksaan tersebut tidak diperlukan. Akan tetapi, keluhan tentang pengulangan menunjukkan adanya celah pemahaman pasien terhadap tujuan dan urutan pelayanan.

Implikasinya, komunikasi tidak sebaiknya bergantung pada kelonggaran waktu atau inisiatif individual dokter dan perawat. Rumah sakit dapat menanamkannya sebagai bagian dari alur kerja melalui protokol komunikasi 4T: (1) tindakan yang akan dilakukan; (2) tujuan atau alasan tindakan, termasuk alasan pemeriksaan diulang; (3) tahap berikutnya; dan (4) taksiran waktu atau kondisi yang dapat menyebabkan keterlambatan. Informasi dapat disampaikan secara



singkat oleh petugas yang paling dekat dengan proses dan diperkuat dengan lembar alur, papan status, video singkat, pesan digital, atau petugas pendamping pasien. Untuk informasi kritis, pemahaman perlu dikonfirmasi dengan *teach-back*, yaitu meminta pasien menjelaskan kembali informasi dengan kata-katanya sendiri [9]. Bukti implementasi menunjukkan bahwa pelatihan, dukungan kepada klinisi, serta pengingat dapat membantu penerapan dan keberlanjutan *teach-back* dalam praktik [10].

Penggunaan ABSA memungkinkan sentimen dikaitkan dengan aspek atau tema tertentu sehingga lebih informatif daripada klasifikasi sentimen keseluruhan [2]. Tinjauan terkini menunjukkan bahwa NLP dapat mengekstraksi tema dari umpan balik tidak terstruktur, tetapi penerapannya masih perlu dihubungkan dengan ukuran mutu yang telah digunakan dan mekanisme agar temuan memengaruhi tindakan [2]. Studi analisis ulasan rumah sakit berbantuan model bahasa menunjukkan bahwa temuan daring dapat membantu manajer membentuk hipotesis untuk investigasi lebih lanjut secara cepat [3]. Oleh karena itu, hasil ABSA RS Mata Makassar perlu ditriangulasi secara rutin dengan Survei Kepuasan Masyarakat, data komplain resmi, *timestamp* perjalanan pasien, jadwal petugas, serta data pembatalan atau kunjungan ulang.

Penelitian ini memiliki beberapa keterbatasan. Ulasan Google Maps bersifat *self-selected*; identitas dan status sebagai pasien tidak dapat diverifikasi; kelompok yang tidak menggunakan platform digital dapat kurang terwakili; dan analisis hanya mencakup satu rumah sakit pada periode tertentu [1]. Pendekatan *rule-based* juga dapat salah menangkap negasi, sarkasme, singkatan, serta konteks lokal atau istilah khusus bidang kesehatan [11]. Penelitian selanjutnya perlu melakukan validasi manual pada sampel berstrata oleh minimal dua penilai serta melaporkan kesepakatan antarpemilai dan kinerja klasifikasi setiap kelas sentimen.

SIMPULAN

Ulasan Google Maps RS Mata Makassar didominasi pengalaman positif, terutama pada *Reliability* dan *Empathy*. Namun, *Responsiveness* merupakan prioritas perbaikan karena memiliki proporsi negatif tertinggi dan menyumbang hampir setengah dari seluruh penetapan negatif. Nilai utama penelitian ini bukan pada klaim bahwa ulasan daring mewakili seluruh pasien, melainkan pada kemampuannya menemukan sinyal proses yang spesifik. Perbaikan yang paling layak diuji adalah transparansi waktu tunggu, protokol komunikasi 4T, pencatatan waktu per tahap, dan uji *one-stop service* berbasis pemetaan perjalanan pasien. NLP-SERVQUAL sebaiknya diintegrasikan dengan survei kepuasan dan data operasional untuk membentuk sistem pemantauan mutu yang berkelanjutan.



REFERENSI

1. Boylan AM, Williams V, Powell J. Online patient feedback: a scoping review and stakeholder consultation to guide health policy. *J Health Serv Res Policy*. 2020;25(2):122–9. doi:10.1177/1355819619870837.
2. Feizollah A, Lin CY, O'Malley L, Thompson W, Listl S, Byrne M. The use of natural language processing to interpret unstructured patient feedback on health services: scoping review. *J Med Internet Res*. 2025;27:e72853. doi:10.2196/72853.
3. Sulaiman MH, Muda N, Abdul Razak F. Analyzing patient complaints in web-based reviews of private hospitals in Selangor, Malaysia, using large language model–assisted content analysis: mixed methods study. *JMIR Form Res*. 2025;9:e69075. doi:10.2196/69075.
4. Parasuraman A, Zeithaml VA, Berry LL. SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality. *J Retailing*. 1988;64(1):12–40.
5. Lau SYZ, Tse MP, Dhalla I, Nayyar D. Understanding online hospital patient reviews: a nationwide cross-sectional analysis of polarity, actionability and themes. *BMJ Open Qual*. 2026;15(3):e004117. doi:10.1136/bmjoc-2025-004117.
6. Zhang H, Ma W, Zhou S, Zhu J, Wang L, Gong K. Effect of waiting time on patient satisfaction in outpatient: an empirical investigation. *Medicine (Baltimore)*. 2023;102(40):e35184. doi:10.1097/MD.00000000000035184.
7. Mbwooge M, Astbury N, Nkumbe HE, Bunce C, Bascaran C. Waiting time and patient satisfaction in a subspecialty eye hospital using a mobile data collection kit: pre-post quality improvement intervention. *JMIRx Med*. 2022;3(3):e34263. doi:10.2196/34263.
8. Lee S, Groß SE, Pfaff H, Dresen A. Waiting time, communication quality, and patient satisfaction: an analysis of moderating influences on the relationship between perceived waiting time and the satisfaction of breast cancer patients during their inpatient stay. *Patient Educ Couns*. 2020;103(4):819–25. doi:10.1016/j.pec.2019.11.018.
9. Shersher V, Haines TP, Sturgiss L, Weller C, Williams C. Definitions and use of the teach-back method in healthcare consultations with patients: a systematic review and thematic synthesis. *Patient Educ Couns*. 2021;104(1):118–29. doi:10.1016/j.pec.2020.07.026.
10. Talevski J, Wong Shee A, Rasmussen B, Kemp G, Beauchamp A. Teach-back: a systematic review of implementation and impacts. *PLoS One*. 2020;15(4):e0231350. doi:10.1371/journal.pone.0231350.
11. Villanueva-Miranda I, Xie Y, Xiao G. Sentiment analysis in public health: a systematic review of the current state, challenges, and future directions. *Front Public Health*. 2025;13:1609749. doi:10.3389/fpubh.2025.1609749.