

## Analisis Sentimen terhadap Rumah Sakit Tarutung Berdasarkan *Google Maps Review* Menggunakan Algoritma *Naïve Bayes* dan Pembobotan Kata *Tf-Idf*

Helmud Laudeweek Panggabean<sup>1</sup>, Farida Gultom<sup>2</sup>, M Bobbi Kurniawan Nasution<sup>3</sup>

<sup>1,2,3</sup>Universitas Efarina

[helmudpgbn@gmail.com](mailto:helmudpgbn@gmail.com)



\*Helmud Laudeweek Panggabean

### Histori Artikel:

Submit: 2025-11-22

Diterima: 2025-12-19

Dipublikasikan: 2025-12-19

### Kata Kunci:

Google Maps Review; Inverse Document Frequency (TF-IDF); Naïve Baye; Rumah Sakit; Term Frequency (TF)

### ABSTRAK

Perkembangan teknologi informasi mendorong masyarakat untuk menyampaikan opini dan pengalaman terhadap layanan publik melalui platform digital, salah satunya Google Maps. Ulasan pengguna pada platform tersebut menjadi sumber data yang penting untuk mengevaluasi kualitas pelayanan rumah sakit. Penelitian ini bertujuan untuk menganalisis sentimen masyarakat terhadap Rumah Sakit Tarutung berdasarkan ulasan Google Maps menggunakan algoritma Naïve Bayes dengan pembobotan kata Term Frequency–Inverse Document Frequency (TF-IDF). Penelitian ini menggunakan pendekatan eksperimental dengan data berupa 250 ulasan pengguna yang dikumpulkan melalui teknik web scraping. Data teks diproses melalui tahapan pra-proses, meliputi case folding, tokenizing, stopword removal, dan stemming, untuk menghasilkan representasi teks yang lebih bersih dan relevan. Selanjutnya, fitur teks dibobotkan menggunakan metode TF-IDF dan diklasifikasikan ke dalam dua kategori sentimen, yaitu positif dan negatif, menggunakan algoritma Naïve Bayes. Evaluasi kinerja model dilakukan menggunakan confusion matrix dengan pembagian data latih sebesar 80% dan data uji sebesar 20%. Hasil penelitian menunjukkan bahwa 70% ulasan memiliki sentimen positif dan 30% memiliki sentimen negatif. Model klasifikasi menghasilkan tingkat akurasi sebesar 85%, dengan nilai presisi 83%, recall 86%, dan F1-score 84%. Hasil ini menunjukkan bahwa algoritma Naïve Bayes berbasis TF-IDF mampu mengklasifikasikan sentimen ulasan Google Maps secara efektif pada data teks berbahasa Indonesia. Temuan penelitian ini dapat dimanfaatkan sebagai bahan evaluasi kualitas pelayanan rumah sakit serta sebagai penerapan analisis sentimen dalam bidang text mining.

Jurnal Pendidikan Sains dan Komputer is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).

### LATAR BELAKANG

Kesehatan adalah sesuatu yang amat penting untuk manusia. Kesehatan merupakan aset paling berharga untuk manusia selain uang dan harta (Rahman, 2021). Dalam menjaga kesehatan negara ambil bagian dalam menjaga dan menjamin kesehatan dari tiap-tiap masyarakatnya. Namun menjaga dan menjamin yang dimaksud di sini bukan berarti keharusan melainkan Negara mengambil peran yang berbeda dari masyarakat. Menjaga pola hidup sehat adalah tugas dari individunya sedangkan kewajiban dan jaminan negara yang dimaksud di sini adalah menyediakan sarana yang baik untuk masyarakat sehingga kesehatan masyarakat tetap terjaga dengan baik. Ada banyak faktor yang akan menjadi penyebab yang mengakibatkan terjadinya gangguan kesehatan pada masyarakat tapi pada penelitian ini penulis akan fokus pada sektor kesehatan di mana negara harus menyediakan fasilitas kesehatan yang memadai untuk masyarakatnya sehingga ketika masyarakat mengalami gangguan kesehatan masyarakat dapat langsung mendapat penanganan yang tepat sehingga dapat kembali pulih dan sehat. (Nainggolan & Sitabuana, Mei 2022).

Perkembangan teknologi informasi telah mendorong perubahan signifikan dalam cara masyarakat



menyampaikan opini dan pengalaman terhadap berbagai layanan publik, termasuk layanan kesehatan. Platform digital seperti Google Maps memungkinkan pengguna untuk memberikan ulasan dan penilaian terhadap rumah sakit berdasarkan pengalaman yang mereka alami secara langsung. Ulasan daring tersebut menjadi sumber informasi yang penting karena mencerminkan tingkat kepuasan dan persepsi masyarakat terhadap kualitas pelayanan rumah sakit (Hu, Zhang, & Pavlou, 2019).

Analisis sentimen merupakan salah satu teknik dalam text mining yang bertujuan untuk mengidentifikasi dan mengklasifikasikan opini atau emosi yang terkandung dalam suatu teks ke dalam kategori tertentu, seperti positif dan negatif (Liu, 2012). Dalam beberapa penelitian sebelumnya, analisis sentimen terbukti mampu memberikan gambaran objektif mengenai kepuasan pengguna terhadap suatu layanan berdasarkan data teks tidak terstruktur (Pang, Lee, & Vaithyanathan, 2002). TF-IDF merupakan metode pembobotan kata dengan menghitung nilai *Term Frequency* dan menghitung banyaknya kemunculan sebuah kata pada koleksi dokumen teks secara keseluruhan. (Mahendra, Murdiansyah, & Lhaksmana, 2023)

Beberapa penelitian terdahulu menunjukkan bahwa kombinasi algoritma Naïve Bayes dan pembobotan TF-IDF memberikan hasil yang baik dalam klasifikasi sentimen ulasan daring, termasuk pada sektor layanan kesehatan (Putra & Wahyudi, 2020; Sari & Lestari, 2019). Namun, penelitian analisis sentimen terhadap ulasan Google Maps pada Rumah Sakit Tarutung masih terbatas. Oleh karena itu, penelitian ini dilakukan untuk menganalisis sentimen ulasan Google Maps terhadap Rumah Sakit Tarutung menggunakan algoritma Naïve Bayes dan pembobotan TF-IDF sebagai upaya memberikan kontribusi ilmiah sekaligus masukan praktis bagi pihak manajemen rumah sakit.

#### STUDI LITERATUR

Analisis sentimen merupakan bagian dari text mining yang bertujuan untuk mengidentifikasi dan mengklasifikasikan opini atau pendapat dalam data teks ke dalam kategori tertentu, seperti positif dan negatif. Feldman (2013) menjelaskan bahwa analisis sentimen banyak digunakan untuk mengekstraksi pengetahuan dari data teks tidak terstruktur yang berasal dari platform digital. Dalam pengolahan teks, representasi fitur memiliki peran penting dalam menentukan kinerja algoritma klasifikasi. Manning, Raghavan, dan Schütze (2008) menyatakan bahwa salah satu metode representasi teks yang paling umum digunakan adalah Term Frequency–Inverse Document Frequency (TF-IDF). Metode ini memberikan bobot lebih tinggi pada kata yang memiliki tingkat kepentingan tinggi dalam suatu dokumen dan relatif jarang muncul pada dokumen lain. Algoritma Naïve Bayes merupakan metode klasifikasi probabilistik yang mengasumsikan independensi antar fitur. Menurut Han, Kamber, dan Pei (2012), Naïve Bayes memiliki kompleksitas komputasi yang rendah serta performa yang stabil, sehingga banyak digunakan dalam klasifikasi teks, termasuk analisis sentimen. Nurhuda, Sihwi, dan Doewes (2013) membuktikan bahwa algoritma Naïve Bayes mampu mengklasifikasikan sentimen teks berbahasa Indonesia dengan tingkat akurasi yang baik.

Penelitian Sari dan Lestari (2019) menunjukkan bahwa analisis sentimen terhadap ulasan layanan kesehatan berbasis media sosial mampu memberikan gambaran yang objektif mengenai kualitas layanan dari sudut pandang pasien. Ulasan daring juga dinilai lebih jujur karena ditulis secara sukarela tanpa intervensi pihak penyedia layanan.

Kombinasi antara pembobotan TF-IDF dan algoritma Naïve Bayes telah banyak digunakan dalam penelitian analisis sentimen. Menurut Aggarwal (2018), penggunaan TF-IDF dalam pengolahan teks mampu meningkatkan kualitas representasi fitur sehingga berdampak positif pada performa algoritma klasifikasi. Penelitian oleh Wahyudi dan Nugroho (2020) menunjukkan bahwa penggunaan Naïve Bayes berbasis TF-IDF pada ulasan Google Maps mampu memberikan hasil klasifikasi sentimen yang cukup akurat.

Berdasarkan penelitian-penelitian tersebut, dapat disimpulkan bahwa algoritma Naïve Bayes dengan pembobotan TF-IDF masih relevan dan efektif digunakan dalam analisis sentimen data teks berbahasa Indonesia. Oleh karena itu, penelitian ini menerapkan metode tersebut sebagai pendekatan komputasi untuk mengklasifikasikan sentimen ulasan Google Maps.

### METODE

Penelitian ini menggunakan pendekatan eksperimental dalam bidang komputasi. Data penelitian berupa 250 ulasan Google Maps Rumah Sakit Tarutung yang dikumpulkan menggunakan teknik web scraping. Tahapan penelitian meliputi: (1) Praproses teks, terdiri dari case folding, tokenizing, stopword removal, dan stemming; (2) Pembobotan kata menggunakan metode TF-IDF; (3) Klasifikasi sentimen menggunakan algoritma Naïve Bayes; (4) Evaluasi model menggunakan confusion matrix. Evaluasi dilakukan untuk mengukur akurasi, presisi, recall, dan F1-score.

### HASIL

Penelitian ini menghasilkan sistem analisis sentimen yang mampu mengklasifikasikan ulasan Google Maps ke dalam sentimen positif dan negatif menggunakan algoritma Naïve Bayes berbasis TF-IDF. Dataset yang digunakan terdiri dari 250 ulasan pengguna Google Maps Rumah Sakit Tarutung yang telah melalui tahapan praproses teks. Tahapan praproses teks berhasil mengubah data ulasan mentah menjadi data yang siap dianalisis. Proses case folding, tokenizing, stopword removal, dan stemming menghasilkan kata-kata dasar yang lebih representatif serta mengurangi noise pada data. Hasil praproses ini berkontribusi terhadap peningkatan kualitas fitur yang digunakan pada tahap klasifikasi. Distribusi Sentimen makah hasil klasifikasi sentimen terhadap 250 ulasan Google Maps ditunjukkan pada Tabel 1.

**Tabel 1. Distribusi Hasil Klasifikasi Sentimen**

Sentimen	Jumlah Ulasan	Persentase
Positif	175	70%
Negatif	75	30%
Total	250	100%

Berdasarkan hasil tersebut, dapat diketahui bahwa sebagian besar ulasan pengguna memiliki sentimen positif, sedangkan sisanya menunjukkan sentimen negative.

Hasil evaluasi kinerja algoritma Naïve Bayes dilakukan menggunakan confusion matrix dengan pembagian data latih sebesar 80% dan data uji sebesar 20%. Parameter evaluasi yang digunakan meliputi akurasi, presisi, recall, dan F1-score. Hasil evaluasi ditunjukkan pada Tabel 2.

**Tabel 2. Hasil Evaluasi Kinerja Model**

Parameter	Nilai
Akurasi	85%
Presisi	83%
Recall	86%
F1-Score	84%

Hasil evaluasi menunjukkan bahwa algoritma Naïve Bayes berbasis TF-IDF mampu mengklasifikasikan sentimen ulasan dengan tingkat akurasi yang cukup baik.

### PEMBAHASAN

Hasil penelitian menunjukkan bahwa algoritma Naïve Bayes berbasis TF-IDF mampu mengklasifikasikan sentimen ulasan Google Maps dengan tingkat akurasi sebesar 85%. Nilai akurasi tersebut menunjukkan bahwa metode yang digunakan cukup efektif dalam mengolah dan mengklasifikasikan data teks berbahasa Indonesia. Hal ini sejalan

dengan penelitian Feldman (2013) yang menyatakan bahwa analisis sentimen berbasis text mining dapat memberikan gambaran yang akurat terhadap opini pengguna apabila didukung oleh metode representasi fitur yang tepat.

Dominasi sentimen positif sebesar 70% menunjukkan bahwa sebagian besar pengguna memberikan penilaian yang baik terhadap objek yang dianalisis. Dari sisi komputasi, hasil ini mengindikasikan bahwa sistem mampu mengenali pola kata-kata positif dan negatif secara cukup konsisten setelah melalui tahapan pra-proses dan pembobotan TF-IDF. Penggunaan TF-IDF berperan penting dalam meningkatkan kualitas fitur dengan memberikan bobot yang lebih besar pada kata-kata yang memiliki makna penting dalam dokumen, sebagaimana dijelaskan oleh Manning, Raghavan, dan Schütze (2008).

Nilai presisi sebesar 83% dan recall sebesar 86% menunjukkan bahwa model memiliki keseimbangan yang cukup baik dalam mengklasifikasikan data uji. Recall yang lebih tinggi dibandingkan presisi menunjukkan bahwa sistem relatif mampu mengenali ulasan yang benar-benar termasuk dalam kelas tertentu, meskipun masih terdapat beberapa kesalahan klasifikasi. Kondisi ini merupakan karakteristik umum dari algoritma Naïve Bayes yang mengasumsikan independensi antar fitur, sebagaimana dikemukakan oleh Han, Kamber, dan Pei (2012).

Hasil penelitian ini juga sejalan dengan penelitian Nurhuda, Sihwi, dan Doewes (2013) serta Wahyudi dan Nugroho (2020) yang menyatakan bahwa algoritma Naïve Bayes berbasis TF-IDF masih relevan digunakan dalam analisis sentimen teks berbahasa Indonesia dengan ukuran dataset menengah. Dengan demikian, penerapan algoritma ini dalam penelitian menunjukkan bahwa pendekatan sederhana namun sistematis masih mampu memberikan performa yang kompetitif.

Secara keseluruhan, pembahasan ini menegaskan bahwa penerapan algoritma Naïve Bayes berbasis TF-IDF dapat menjadi solusi yang efektif dalam analisis sentimen berbasis teks. Penelitian ini memberikan kontribusi pada bidang pendidikan sains dan komputer, khususnya dalam penerapan konsep text mining dan klasifikasi teks pada data nyata yang bersumber dari platform digital.

### KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa algoritma Naïve Bayes berbasis pembobotan TF-IDF mampu digunakan secara efektif untuk analisis sentimen ulasan Google Maps. Penerapan tahapan pra-proses teks, pembobotan fitur, dan klasifikasi probabilistik menghasilkan sistem yang mampu mengklasifikasikan sentimen ulasan dengan akurasi sebesar 85%. Hasil analisis sentimen menunjukkan bahwa mayoritas ulasan pengguna memiliki sentimen positif, yaitu sebesar 70%, sedangkan 30% ulasan memiliki sentimen negatif.

Temuan ini menunjukkan bahwa pendekatan text mining dapat dimanfaatkan untuk mengolah data teks tidak terstruktur dan menghasilkan informasi yang berguna secara komputasional. Secara keseluruhan, penelitian ini memberikan kontribusi dalam bidang pendidikan sains dan komputer melalui penerapan algoritma klasifikasi teks pada data nyata. Penelitian selanjutnya disarankan untuk menggunakan dataset yang lebih besar, menambahkan kategori sentimen netral, serta membandingkan kinerja Naïve Bayes dengan algoritma klasifikasi lainnya guna meningkatkan akurasi sistem.

### SARAN

Berdasarkan hasil penelitian yang telah dilakukan, penelitian selanjutnya disarankan untuk menggunakan jumlah data ulasan yang lebih besar agar model klasifikasi dapat dilatih secara lebih optimal. Selain itu, penambahan kelas sentimen netral diharapkan dapat memberikan hasil analisis yang lebih komprehensif. Penelitian berikutnya juga dapat membandingkan algoritma Naïve Bayes dengan metode klasifikasi lainnya, seperti Support Vector Machine atau Random Forest, guna mengetahui metode yang memiliki kinerja terbaik. Penggunaan teknik pembobotan atau representasi fitur lainnya, seperti word embedding, juga dapat dipertimbangkan untuk meningkatkan akurasi analisis sentimen.

### UCAPAN TERIMA KASIH

Dalam penyusunan jurnal ini penulis menyampaikan terimakasih pada:



1. Bapak DR. J.R Saragih SelakuBapa Pendiri Yayasan Universitas Efarina
2. Kepada Ibu Farida Gultom S, Kom.,M.Kom Selaku Dekan Fakultas Teknok Universitas Efarina
3. Kepada Bapak M. Bobbi Nasution
4. Kepada keluarga dan juga teman-teman yang sudah membantu

#### REFERENSI

- Asrumi, Suharijadi, D., Setiari, A. D., & Wulanda, D. P. (2023). Analisis Sentimen dan Penggali Opini.
- Aggarwal, C. C. (2018). *Machine learning for text*. Springer.
- Feldman, R. (2013). Techniques and applications for sentiment analysis. *Communications of the ACM*, 56(4), 82–89. <https://doi.org/10.1145/2436256.2436274>
- Han, J., Kamber, M., & Pei, J. (2012). *Data mining: Concepts and techniques* (3rd ed.). Morgan Kaufmann.
- Hu, N., Zhang, J., & Pavlou, P. A. (2019). Overcoming the J-shaped distribution of product reviews. *Communications of the ACM*, 62(10), 96–104. (f)
- Liu, B. (2012). *Sentiment analysis and opinion mining*. Morgan & Claypool.
- Mahendra, M. H., Murdiansyah, D. T., & Lhaksana, M. K. (2023). Analisis Sentimen Tweet COVID-19 menggunakan K-Nearest Neighbors dengan TF-IDF dan Ekstraksi Fitur CountVectorizer.
- Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to information retrieval*. Cambridge University Press.
- Nainggolan, V., & Sitabuana, H. T. (Mei 2022). Jaminan Kesehatan Bagi Rakyat Indonesia Menurut Hukum Kesehatan. (f)
- Nurhuda, F., Sihwi, S. W., & Doewes, A. (2013). Analisis sentimen masyarakat terhadap calon presiden Indonesia 2014 berdasarkan opini dari Twitter menggunakan metode Naïve Bayes Classifier. *Jurnal Teknologi & Informasi ITSmart*, 2(2), 35–42.
- Pang, B., Lee, L., & Vaithyanathan, S. (2002). Thumbs up? Sentiment classification using machine learning techniques. *Proceedings of EMNLP*, 79–86.
- Putra, A. R., & Wahyudi, M. (2020). Penerapan TF-IDF pada klasifikasi teks ulasan. *Jurnal Ilmu Komputer*, 8(1), 12–20
- Rahman, H. (2021). Klasifikasi Sentimen Masyarakat Terhadap Layanan Badan Penyelenggara Jaminan Sosial (Bpjs) Kesehatan Di Twitter Menggunakan Metode K-Nearest Neighbor. (f)
- Sari, D. P., & Lestari, R. (2019). Analisis sentimen layanan kesehatan berbasis media sosial. *Jurnal Sistem Informasi*, 11(2), 101–110