

Optimasi Metode Naïve Bayes Classifier dalam Analisis Sentimen RSUD Siti Fatimah Palembang

¹Ardiansyah, ²Kurniawan

^{1,2}Universitas Bina Darma, Indonesia

¹ardi48298@gmail.com; ²kurniawan@binadarma.ac.id

Article Info

Article history:

Received, xxx xx xxxx

Revised, xxx xx xxxx

Accepted, xxx xx xxx

Kata Kunci:

analisis_sentimen

naïve_bayes_classifier

web_scraping

tf-idf

rsud_siti_fatimah

ABSTRAK

Rumah Sakit Umum Daerah Siti Fatimah Palembang (RSUD) merupakan institusi kesehatan penting yang dioperasikan pemerintah yang telah mengumpulkan banyak evaluasi dari masyarakat mengenai pemberian layanannya. Tujuan utama dari penelitian ini adalah untuk melakukan analisis sentimen publik yang diarahkan terhadap RSUD Siti Fatimah, menggunakan metodologi Naïve Bayes Classifier. Pendekatan analitis ini digunakan untuk secara sistematis mengkategorikan ulasan ke dalam sentimen positif dan negatif. Data yang berkaitan dengan ulasan diperoleh melalui teknik pengikisan web dari Google Maps, diikuti oleh serangkaian prosedur prapemrosesan teks, yang mencakup sanitasi teks, tokenisasi, dan penerapan TF-IDF untuk pembobotan. Hasil penyelidikan ini menjelaskan distribusi sentimen, sehingga memberikan pemahaman yang komprehensif tentang persepsi publik mengenai layanan yang diberikan oleh Rumah Sakit Siti Fatimah. Temuan penelitian ini diharapkan akan memberi kontribusi yang tak ternilai bagi administrasi rumah sakit dalam upaya mereka untuk meningkatkan kualitas layanan.

ABSTRACT

Siti Fatimah General Hospital (RSUD) in Palembang is an important government-operated healthcare institution that has gathered numerous evaluations from the public regarding its service delivery. The primary objective of this research is to conduct a public sentiment analysis directed at RSUD Siti Fatimah, using the Naïve Bayes Classifier methodology. This analytical approach is employed to systematically categorize reviews into positive and negative sentiments. Review data was obtained through web scraping techniques from Google Maps, followed by a series of text preprocessing procedures, including text sanitization, tokenization, and the application of TF-IDF for weighting. The findings of this investigation outline the sentiment distribution, thus providing a comprehensive understanding of public perceptions of the services provided by Siti Fatimah Hospital. The results of this research are expected to contribute invaluable insights to the hospital administration in their efforts to improve service quality.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-nc-nd/4.0/) license.



Penulis Korespondensi:

Jojo,

Program Studi Informatika,

Universitas Komputer Indonesia,

Email: jojo@gmail.com

1. PENDAHULUAN

Rumah Sakit Umum Daerah (RSUD) berfungsi sebagai institusi kesehatan di bawah kepemilikan pemerintah daerah, yang meliputi Rumah Sakit Siti Fatimah yang berlokasi di Palembang. Sebagai penyedia layanan kesehatan terkemuka di Sumatera Selatan, Rumah Sakit Siti Fatimah memberikan kontribusi signifikan terhadap penyediaan layanan kesehatan kepada penduduk sekitarnya. Evaluasi publik mengenai layanan yang diberikan oleh rumah sakit ini sering disampaikan melalui berbagai platform, termasuk Google Maps. Evaluasi semacam itu sangat penting karena menawarkan wawasan komprehensif tentang persepsi

masyarakat tentang layanan rumah sakit. Sentimen publik yang berlaku mengenai layanan rumah sakit dapat berfungsi sebagai indikator tingkat kepuasan atau ketidakpuasan, yang pada gilirannya mempengaruhi citra publik lembaga dan kepercayaan masyarakat.

Tujuan dari penelitian ini adalah untuk melakukan analisis sentimen tinjauan publik yang berkaitan dengan RSUD Siti Fatimah dengan menggunakan metodologi Naïve Bayes Classifier, algoritma pembelajaran mesin yang digunakan untuk klasifikasi teks. Pendekatan Naïve Bayes diakui sebagai salah satu teknik unggulan dalam analisis sentimen karena keterusterangannya dan kemampuannya untuk mengelola kumpulan data berbasis teks yang luas dengan mahir[1]. Dalam ranah analisis sentimen, teknik ini digunakan untuk mengkategorikan teks ulasan menjadi dua klasifikasi berbeda: positif dan negatif.

Analisis Sentimen merupakan proses sistematis memahami, mengekstraksi, dan mengkategorikan pendapat atau ekspresi emosional dari teks yang dibuat pengguna, seperti ulasan atau komentar. Tujuan utama analisis sentimen adalah untuk memastikan apakah teks yang dianalisis menyampaikan sentimen positif, negatif, atau netral. Dalam konteks penelitian ini, analisis sentimen digunakan untuk menjelaskan persepsi masyarakat mengenai layanan yang diberikan oleh Rumah Sakit Siti Fatimah, yang berfungsi sebagai umpan balik penting bagi institusi kesehatan yang berusaha meningkatkan kualitas layanan[2].

Naïve Bayes Classifier (NBC) beroperasi berdasarkan Teorema Bayes untuk meramalkan klasifikasi teks berdasarkan probabilitas terkait. Algoritma ini sering digunakan dalam analisis sentimen karena kemampuannya dalam mengklasifikasikan data teks, ditambah dengan kemampuannya dalam memproses kumpulan data substansial. Sejumlah penelitian sebelumnya, seperti yang dilakukan oleh Rahman et al. (2024), menunjukkan bahwa Naïve Bayes menunjukkan kinerja yang terpuji dalam klasifikasi teks, terutama dalam menilai kelayakan pelanggan dalam domain perbankan[3]. Dalam studi saat ini, Naïve Bayes disandingkan dengan metodologi Random Forest, menghasilkan hasil yang menunjukkan Naïve Bayes lebih unggul dalam mengklasifikasikan data berdasarkan probabilitas.

Penelitian lebih lanjut yang dilakukan oleh Khofifah dkk. (2022) menguatkan kemanjuran Naïve Bayes dalam melakukan analisis sentimen pada ulasan pengguna mengenai tempat wisata di Google Maps[4]. Temuan penyelidikan ini mengungkapkan bahwa Naïve Bayes mencapai akurasi yang terpuji dalam klasifikasi teks berorientasi sentimen, menjadikannya metodologi yang cocok untuk menganalisis ulasan layanan di RSUD Siti Fatimah.

Investigasi oleh Maulana et al. (2023), dalam menjalankan analisis sentimen aplikasi MyPertamina menggunakan algoritma NBC, menunjukkan bahwa metodologi ini efektif dalam menafsirkan sentimen ulasan pengguna[5]. Selanjutnya, penelitian oleh Derwis et al. (2021), yang menerapkan teknik Naïve Bayes untuk analisis sentimen data Twitter, juga menunjukkan bahwa algoritma ini dapat berperan dalam membedakan opini publik mengenai layanan publik[6].

Mengambil wawasan dari berbagai penelitian sebelumnya, penerapan Naïve Bayes Classifier dalam analisis sentimen di Rumah Sakit Siti Fatimah diantisipasi untuk menghasilkan hasil yang tepat, yang pada akhirnya dapat berfungsi sebagai sumber daya yang berharga bagi administrasi rumah sakit dalam upaya mereka untuk meningkatkan kualitas layanan kesehatan.

2. METODE PENELITIAN

Metode penelitian menggunakan metodologi Naïve Bayes Classifier yang dipergunakan untuk melakukan analisis sentimen komprehensif dari ulasan yang dikumpulkan secara sistematis dari platform Google Maps, menggunakan teknik scraping yang dirancang untuk mengekstrak data tekstual yang relevan.

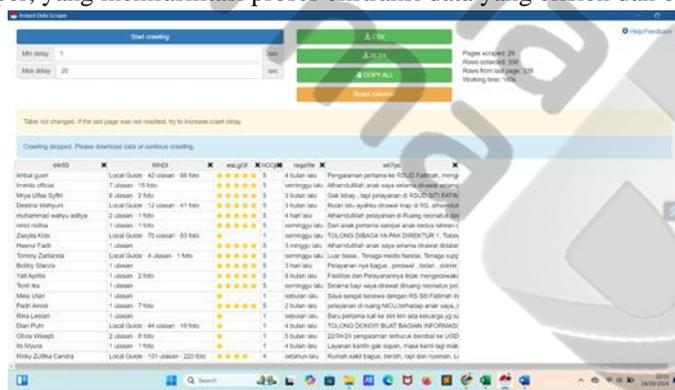


Gambar 1 Langkah metodologis

Langkah-langkah metodologis yang dilakukan selama penelitian ini digambarkan sebagai berikut:

1. Pengumpulan Data

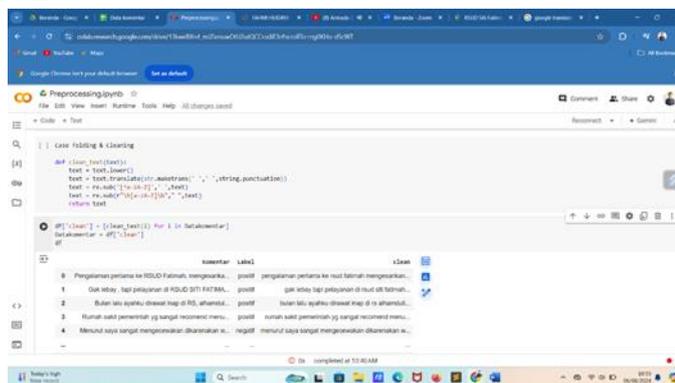
Tahap awal penelitian melibatkan pengumpulan data sistematis yang berasal dari ulasan buatan pengguna yang tersedia di platform Google Maps, dicapai melalui pemanfaatan ekstensi browser Instant Data Scraper, yang memfasilitasi proses ekstraksi data yang efisien dan otomatis.



Gambar 2 Scraping data

2. Teks preprocessing

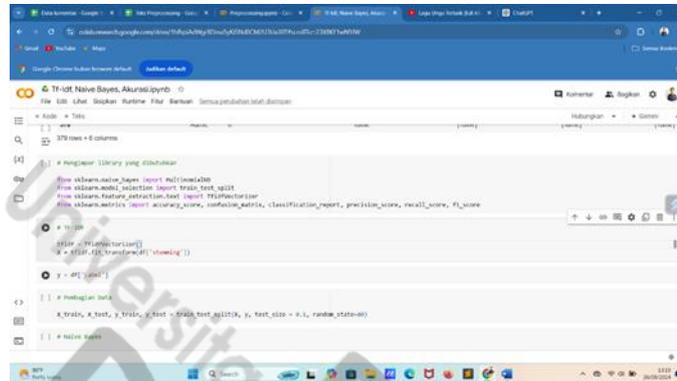
Tahap selanjutnya penelitian ini mencakup serangkaian proses diantaranya mencakup case folding, cleaning, tokenization, stopwords removal, dan stemming yang bertujuan untuk membersihkan dan menyederhanakan teks.



Gambar 3 Hasil case folding dan cleaning

3. Pembobotan Kata

Fase berikutnya melibatkan penerapan pendekatan Term Frequency-Inverse Document Frequency (TF-IDF), yang berfungsi untuk menetapkan bobot yang dapat diukur untuk setiap kata individu dalam kumpulan data, sehingga memungkinkan pemahaman bernuansa tentang pentingnya setiap istilah relatif terhadap keseluruhan korpus.



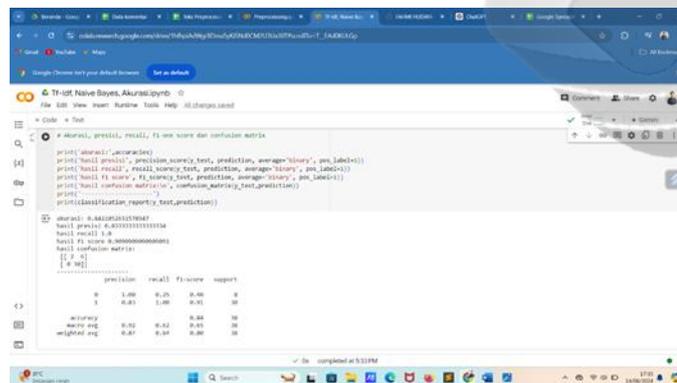
Gambar 4 Proses TF-IDF

4. Klasifikasi

Dalam langkah penting ini, model klasifikasi Naïve Bayes digunakan untuk secara sistematis mengkategorikan sentimen yang diungkapkan dalam ulasan ke dalam klasifikasi positif dan negatif yang berbeda, sehingga memfasilitasi interpretasi terstruktur dari nada emosional yang disampaikan oleh pengguna.

5. Evaluasi

Tahap terakhir, kinerja model klasifikasi dinilai secara ketat melalui penerapan berbagai metrik evaluatif, termasuk akurasi, presisi, recall, dan skor F1, yang secara kolektif memberikan gambaran komprehensif tentang efektivitas model sehingga secara akurat membedakan sentimen yang ada dalam data ulasan.



Gambar 5 Hasil proses evaluasi model

3. HASIL DAN ANALISIS

Hasil dari penelitian ini menunjukkan bahwa mayoritas ulasan yang diterima oleh RSUD Siti Fatimah Palembang bersifat positif, mengindikasikan adanya tingkat kepuasan yang cukup tinggi dari masyarakat terkait pelayanan yang diberikan oleh rumah sakit. Ulasan-ulasan ini sering kali memuji aspek-aspek seperti fasilitas yang memadai, kebersihan rumah sakit, serta profesionalisme tenaga medis. Hal ini menunjukkan bahwa RSUD Siti Fatimah telah mampu memenuhi ekspektasi pasien dalam beberapa aspek penting, yang secara tidak langsung memperkuat citra positif rumah sakit di mata masyarakat.

Namun demikian, ada sejumlah ulasan negatif yang memfokuskan pada beberapa kelemahan dalam pelayanan. Ulasan negatif ini umumnya berkaitan dengan waktu tunggu yang lama, baik untuk proses administrasi, pelayanan medis, serta sikap tenaga medis yang dinilai kurang ramah dan tanggap terhadap

keluhan pasien. Masalah ini menjadi perhatian penting, karena persepsi negatif dapat mempengaruhi kepuasan pasien secara keseluruhan dan menurunkan reputasi institusi kesehatan di mata masyarakat.

Metode Naïve Bayes Classifier digunakan untuk mengklasifikasi ulasan dalam kategori positif dan negatif. Naïve Bayes terbukti sangat efektif dalam mengidentifikasi pola dalam teks ulasan, yang kemudian digunakan untuk mengukur sentimen secara keseluruhan. Penggunaan metode ini melibatkan beberapa tahap penting, seperti text preprocessing mencakup pembersihan data, tokenisasi, dan pembobotan menggunakan TF-IDF. Proses ini membantu memastikan data yang digunakan dalam klasifikasi relevan dan bersih dari noise.

Setelah tahap preprocessing, model Naïve Bayes dilatih menggunakan data ulasan yang telah dilabeli, dan hasil klasifikasi dievaluasi menggunakan confusion matrix. Evaluasi menunjukkan tingkat akurasi model cukup tinggi, dengan nilai precision, recall, dan F1-score yang memadai untuk kedua kategori sentimen. Tingkat akurasi ini menunjukkan bahwa Naïve Bayes mampu mengklasifikasikan ulasan dengan baik, terutama dalam membedakan ulasan positif dan negatif berdasarkan pola kata yang muncul.

Dalam pengukuran precision dan recall, analisis menunjukkan bahwa model ini lebih baik dalam mengidentifikasi ulasan positif dibandingkan ulasan negatif. Hal ini dapat diakibatkan oleh dominasi ulasan positif dalam dataset, yang menyebabkan bias dalam pelatihan model. Meski demikian, model masih mampu mendeteksi sejumlah ulasan negatif dengan cukup baik, yang menjadi dasar bagi identifikasi aspek-aspek layanan yang perlu ditingkatkan.

Tabel 1 Hasil nilai precision, recall dan f-1 score

Jenis Klasifikasi	Precision	Recall	F-1 Score
Positif	0,83	1,00	0,91
Negatif	1,00	0,25	0,40

Hasil confusion matrix juga menunjukkan bahwa terdapat beberapa false positives dan false negatives, di mana beberapa ulasan yang sebenarnya positif terklasifikasikan sebagai negatif, dan sebaliknya. Ini menunjukkan adanya beberapa kesalahan dalam interpretasi model terhadap ulasan yang mungkin bersifat ambigu atau tidak jelas secara konteks. Ulasan seperti ini sering kali mengandung bahasa yang bercampur antara positif dan negatif, yang membuat klasifikasi menjadi lebih kompleks.

Secara keseluruhan, hasil analisis ini memberikan gambaran yang komprehensif tentang persepsi masyarakat terhadap RSUD. Temuan ini memberikan wawasan penting bagi manajemen rumah sakit untuk memahami kekuatan dan kelemahan layanan, serta area yang perlu diperbaiki, seperti efisiensi waktu tunggu dan sikap tenaga medis. Dengan melakukan perbaikan tersebut, RSUD diharapkan dapat meningkatkan kualitas pelayanan yang lebih baik dan responsif terhadap kebutuhan masyarakat.

Selain itu, temuan ini juga menunjukkan potensi metode Naïve Bayes Classifier dalam analisis sentimen di bidang pelayanan kesehatan. Keberhasilan metode ini dalam mengklasifikasikan ulasan publik berdasarkan sentimen positif dan negatif memperlihatkan bahwa teknik ini dapat diterapkan di berbagai sektor layanan publik lainnya untuk mendapatkan pemahaman yang lebih baik mengenai persepsi masyarakat.

4. KESIMPULAN

Penelitian ini berhasil melakukan analisis sentimen terhadap ulasan masyarakat mengenai RSUD Siti Fatimah Palembang menggunakan metode Naïve Bayes Classifier. Hasilnya memberikan gambaran yang jelas mengenai persepsi masyarakat terhadap layanan rumah sakit ini. Temuan ini diharapkan dapat menjadi bahan pertimbangan bagi manajemen rumah sakit dalam meningkatkan kualitas pelayanan, serta menjadi panduan untuk penelitian lanjutan dalam analisis sentimen di bidang layanan kesehatan.

UCAPAN TERIMA KASIH

Penulis mengucapkan terimakasih kepada segenap pihak yang telah memberi bimbingan dan dukungan kepada penulis sehingga tulisan ini bisa selesai dan bermanfaat.

REFERENSI

- [1] T. Dzulkarnain, D. E. Ratnawati, and B. Rahayudi, "Penggunaan Metode Naïve Bayes Classifier pada Analisis Sentimen Penilaian Masyarakat Terhadap Pelayanan Rumah Sakit di Malang," *J. Teknol. Inf. Dan Ilmu Komput.*, vol. 10(7), pp. 1453–1460, 2023.
- [2] P. A. Permatasari, L. Linawati, and L. Jasa, "Survei Tentang Analisis Sentimen Pada Media Sosial," *Maj. Ilm. Teknol. Elektro*, vol. 20(2), p. 177, 2021, doi: <https://doi.org/10.24843/mite.2021.v20i02.p01>.

- [3] A. Rahman, A. S. Wati, A. Marina, N. I. H. Kunio, and N. Jumrituniisah, "Application of Data Mining for Classification of Customer Eligibility at XYZ Bank in Credit Agreements Using Naives Bayes and Random Forest Methods," in *Proceedings of the 7th FIRST 2023 International Conference on Global Innovations (FIRST-ESCSI 2023)*, vol. 232, N. L. Husni, W. Caesarendra, M. Aznury, L. Novianti, and D. Stiawan, Eds., in *Advances in Engineering Research*, vol. 232. , Dordrecht: Atlantis Press International BV, 2024, pp. 563–569. doi: 10.2991/978-94-6463-386-3_57.
- [4] W. Khofifah, D. N. Rahayu, and A. M. Yusuf, "Analisis Sentimen Menggunakan Naive Bayes Untuk Melihat Review Masyarakat Terhadap Tempat Wisata Pantai Di Kabupaten Karawang Pada Ulasan Google Maps," *J. Interkom*, vol. 16(4), pp. 28–38, 2022.
- [5] R. Maulana, A. Voutama, and T. Ridwan, "Analisis Sentimen Ulasan Aplikasi MyPertamina pada Google Play Store menggunakan Algoritma NBC," *J. Teknol. Terpadu*, vol. 9(1), pp. 42–48, 2023.
- [6] D. Darwis, N. Siskawati, and Z. Abidin, "Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Review Data Twitter BMKG Nasional," *J. Tekno Kompak*, vol. 15(1), p. 131, 2021.
- [7] Dzulkarnain, T., Ratnawati, D. E., & Rahayudi, B. (2023). Penggunaan Metode Naive Bayes Classifier pada Analisis Sentimen Penilaian Masyarakat Terhadap Pelayanan Rumah Sakit di Malang. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 10(7), 1453–1460. <https://doi.org/10.25126/jtiik.1077979>
- [8] Huda, F. (2023). *Analisis Sentimen Terhadap Aplikasi Sapawarga-Jabar Super Apps Pada Ulasan Google Play Store*. 34.
- [9] Pemula, P. (2023). *Skema: Penelitian Pemula. April. Analisis Sentimen Pada CHATGPT Menggunakan Algoritma Transformers*
- [10] Li, B. A. B., & Pustaka, T. (2002). *BAB II Tinjauan Pustaka BAB II TINJAUAN PUSTAKA 2.1*. 1–64.
- [11] Febriana Sulistya Pratiwi. (2022). No, 8.5.2017. <https://dataindonesia.id/sektor-riil/detail/angka-konsumsi-ikan-ri-naik-jadi-5648-kgkapita-pada-2022>
- [12] Zahra, A. (2024). *Analisis Sentimen Masyarakat Terhadap Kualitas Layanan Rumah Sakit Indonesia Menggunakan Algoritma Text Mining dan TF-IDF*. 7(1), 34–42. <https://doi.org/10.30865/komik.v6i1.8036>



Letter of Acceptance

Handwritten signature and date: 9/10/2024

1Ardiansyah, 2Kurniawan
1,2Universitas Bina Darma, Indonesia

Kepada Penulis,

Redaksi Journal Scientific and Applied Informatics dengan ini menyatakan bahwa artikel yang telah disubmit berjudul "Optimasi Metode Naive Bayes Classifier dalam Analisis Sentimen RSUD Siti Fatimah Palembang" diterima sebagai salah satu artikel yang akan diterbitkan pada Volume 7 Nomor 3 Bulan November 2024.

Bengkulu, September 2024
Editor in Chief,



Erwin Dwika Putra, M.Kom