

# Analisis Sentimen Penilaian Masyarakat Terhadap Rumah Sakit Berdasarkan Ulasan Pada Google Maps Menggunakan Naive Bayes

Nadhir Fachrul Rozam<sup>1</sup>, Tika Novita Sari<sup>2</sup>

Program Studi Teknologi Informasi, Universitas Negeri Yogyakarta

Jl. Colombo No.1, Daerah Istimewa Yogyakarta

<sup>1</sup>nadhirfachrulrozam@uny.ac.id, <sup>2</sup>tikanovitasari@uny.ac.id

**Abstract** - Hospitals are vital facilities in the lives of society, where the level of public satisfaction with hospital services is one of the key indicators of hospital performance. One direct assessment from the public can be observed through reviews on Google Maps. The public reviews regarding the services of Muhammadiyah Lamongan Hospital serve as the object of this research. This study aims to analyze the sentiment of public reviews toward the services provided by Muhammadiyah Lamongan Hospital using data from Google Maps. Sentiment analysis was conducted to classify the reviews into two categories: positive and negative, providing an overview of public perception of hospital services. The data collected revealed that 95% of the reviews (3989 data) fell into the positive category, while 5% of the reviews (172 data) were classified as negative. These findings indicate that the majority of visitors have a positive perception of the hospital's services. To build the sentiment analysis system, Naive Bayes algorithm was applied using three model variants: ComplementNB, MultinomialNB, and BernoulliNB. Before the training process, the review data underwent a pre-processing stage consisting of stemming, stopword removal, emoticon normalization, and TF-IDF to prepare the data. The results showed that the BernoulliNB model achieved the best performance with an accuracy of 84.21%, followed by ComplementNB with an accuracy of 80.26%, and MultinomialNB with an accuracy of 72.37%. This research contributes to improving hospital management and services by providing real-time sentiment analysis as feedback that can be used for evaluating and enhancing service quality.

**Keywords** - Naïve Bayes, TF-IDF, Sentiment Analysis, Google Maps, and Hospital.

**Abstrak** - Rumah Sakit menjadi salah satu fasilitas vital dalam kehidupan masyarakat. Tingkat kepuasan masyarakat terhadap pelayanan rumah sakit menjadi salah satu indikator untuk kinerja rumah sakit didalam masyarakat. Salah satu penilaian langsung dari masyarakat dapat dilihat dari ulasan pada Google Maps. Ulasan masyarakat tentang pelayanan Rumah Sakit Muhammadiyah Lamongan menjadi objek penelitian penulis disini. Penelitian ini bertujuan untuk menganalisis sentimen ulasan masyarakat terhadap pelayanan Rumah Sakit Muhammadiyah Lamongan menggunakan data dari Google Maps. Analisis sentimen dilakukan untuk mengklasifikasikan ulasan ke dalam dua kategori, yaitu positif dan negatif, sehingga dapat memberikan gambaran mengenai persepsi masyarakat terhadap pelayanan rumah sakit. Dari data yang didapatkan menunjukkan bahwa 95% ulasan (3989 data) termasuk dalam kategori positif, sementara 5% ulasan (172 data) termasuk dalam kategori negatif. Hasil ini menggambarkan bahwa mayoritas pengunjung memiliki persepsi positif terhadap pelayanan rumah sakit. Selanjutnya untuk membuat sistem analisis sentimen, metode yang digunakan adalah algoritma Naive Bayes dengan tiga varian model, yaitu ComplementNB, MultinomialNB, dan BernoulliNB. Sebelum proses pelatihan, data ulasan melalui tahap pre-processing berupa stemming, penghapusan stopword, normalisasi emoticon, dan pembobotan TF-IDF untuk mempersiapkan data dalam format numerik. Hasil penelitian menunjukkan bahwa BernoulliNB memberikan performa terbaik dengan akurasi sebesar 84,21%, diikuti oleh ComplementNB dengan akurasi 80,26%, dan MultinomialNB dengan akurasi 72,37%. Penelitian ini berkontribusi pada peningkatan pelayanan dan manajemen rumah sakit dengan menyediakan analisis sentimen secara real-time sebagai umpan balik yang dapat digunakan untuk evaluasi dan perbaikan layanan.

**Kata Kunci** - Naïve Bayes, TF-IDF, Analisis Sentimen, Google Maps, dan Rumah Sakit.

## I. PENDAHULUAN

Rumah sakit merupakan salah satu elemen penting dalam sistem kesehatan yang memiliki peran krusial dalam memberikan pelayanan medis kepada masyarakat. Kualitas pelayanan yang diberikan oleh rumah sakit sering kali menjadi tolok ukur kepuasan pasien dan masyarakat umum [1]–[3]. Di era digital saat ini, banyak individu yang mengekspresikan pengalaman dan penilaian mereka terhadap pelayanan

rumah sakit melalui platform daring, salah satunya adalah Google Maps [4]. Ulasan yang diberikan oleh pengguna di Google Maps tidak hanya mencakup rating numerik, tetapi juga opini verbal yang mengandung sentimen positif, negatif, atau netral.

Analisis sentimen merupakan salah satu metode yang dapat digunakan untuk memahami perasaan dan opini masyarakat terhadap suatu objek [5], dalam hal ini rumah sakit. Dengan menggunakan teknik ini, kita dapat menggali informasi berharga dari ulasan

pengguna untuk mengidentifikasi kekuatan dan kelemahan dalam pelayanan yang diberikan oleh rumah sakit. Hal ini memungkinkan pihak manajemen rumah sakit untuk melakukan perbaikan dan peningkatan kualitas layanan berdasarkan masukan langsung dari pasien dan masyarakat.

Naive Bayes adalah salah satu algoritma klasifikasi yang sering digunakan dalam analisis sentimen karena kesederhanaannya dan kemampuannya dalam menangani data teks yang bersifat probabilistik [6]. Algoritma ini bekerja dengan mengklasifikasikan teks berdasarkan probabilitas yang dihitung dari kemunculan kata-kata tertentu dalam teks yang telah diberi label sentimen sebelumnya [7].

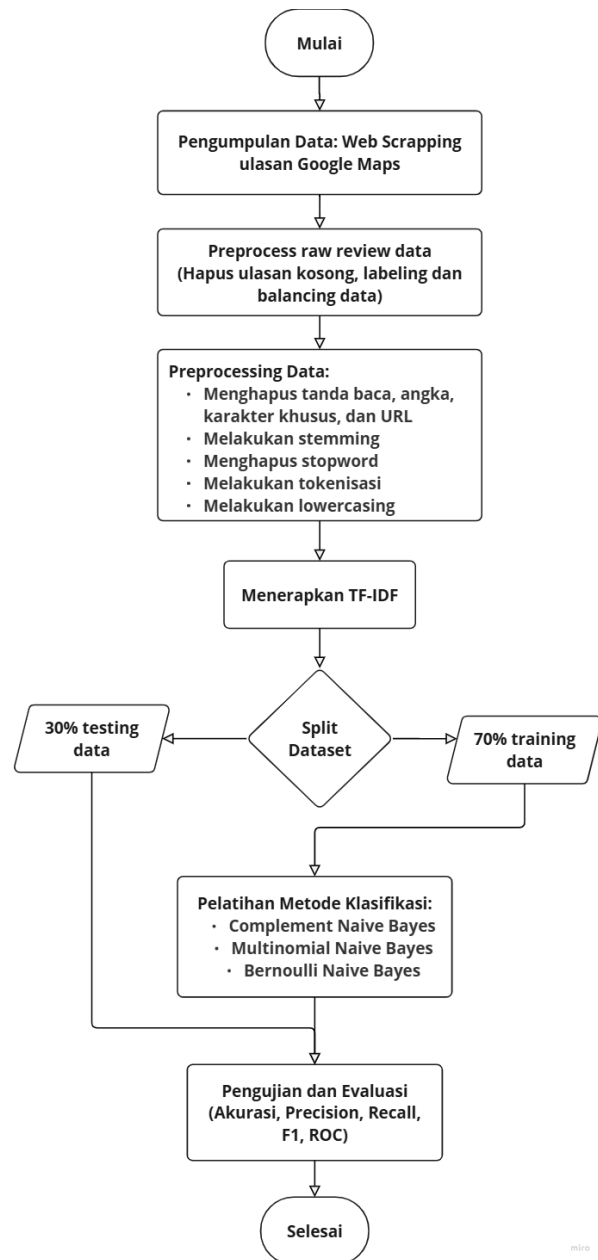
Penelitian terkait analisis sentimen terhadap layanan publik, termasuk rumah sakit, telah banyak dilakukan sebelumnya menggunakan berbagai metode klasifikasi teks, seperti Naive Bayes [2], [7], [8], Support Vector Machine (SVM) [1], dan model berbasis deep learning, seperti RNN [9] dan CNN [10]. Namun, penelitian-penelitian tersebut masih menghadapi keterbatasan, seperti ketidakseimbangan data, rendahnya representasi opini pada platform tertentu, serta minimnya eksplorasi performa varian Naive Bayes yang lebih detail untuk menangani data ulasan. Selain itu, sebagian besar studi berfokus pada analisis sentimen berbasis media sosial, seperti Twitter atau Facebook, dan masih jarang yang memanfaatkan ulasan dari platform Google Maps, yang sebenarnya memiliki tingkat relevansi tinggi karena menampilkan opini langsung dari pengunjung fasilitas tertentu [3]. Dengan demikian, belum banyak penelitian yang secara khusus menganalisis ulasan layanan rumah sakit yang bersumber dari Google Maps dan membandingkan performa beberapa varian Naive Bayes, seperti ComplementNB, MultinomialNB, dan BernoulliNB, dalam menangani klasifikasi sentimen.

Dalam penelitian ini, dilakukan analisis sentimen terhadap ulasan masyarakat terhadap salah satu rumah sakit di Google Maps dengan menggunakan metode Naive Bayes. Penelitian ini bertujuan untuk memahami bagaimana persepsi masyarakat terhadap kualitas layanan rumah sakit, serta untuk mengeksplorasi sejauh mana ulasan daring dapat dijadikan sebagai indikator penilaian publik. Dengan demikian, penelitian ini diharapkan dapat memberikan wawasan yang lebih mendalam bagi pengelola rumah sakit untuk terus meningkatkan kualitas pelayanan kesehatan mereka, serta bagi peneliti dalam bidang analisis data teks dan pemrosesan bahasa alami.

## II. METODE PENELITIAN

Bab ini menjelaskan langkah-langkah yang dilakukan dalam penelitian untuk menganalisis sentimen ulasan masyarakat terhadap Rumah Sakit Muhammadiyah Lamongan menggunakan metode Naive Bayes. Metode ini terdiri dari empat tahapan utama, yaitu Pengumpulan Data, Preprocessing Data,

Pelatihan Metode Klasifikasi, dan Pengujian dan Evaluasi. Keempat tahapan ini dilakukan secara berurutan untuk memastikan bahwa analisis sentimen dilakukan secara sistematis dan menghasilkan hasil yang akurat. Untuk lebih jelasnya dapat dilihat pada Gambar 1 untuk alur penelitian.



Gambar 1. Flowchart Penelitian

### A. Mengumpulkan Data

Data untuk penelitian ini diperoleh dari ulasan pengguna pada platform Google Maps terkait Rumah Sakit Muhammadiyah Lamongan. Proses pengumpulan data dilakukan dengan menggunakan metode *web scraping*, yang mematuhi kebijakan privasi dan aturan penggunaan platform. Untuk melaksanakan *web scraping*, digunakan beberapa tools

dan library Python, seperti Botasaurus [11], untuk mengambil data dari Google Maps secara otomatis.

Dataset yang diperoleh terdiri dari 4777 ulasan, dengan rincian sebagai berikut:

1. 616 ulasan kosong dihapus karena tidak relevan.
2. Dataset akhir berjumlah 4161 ulasan, yang digunakan untuk klasifikasi sentimen biner.

Dataset ini terdiri dari 172 ulasan negatif dan 3989 ulasan positif, yang menunjukkan bahwa data bersifat tidak seimbang (imbalanced). Untuk mengatasi ketidakseimbangan ini, dilakukan pemilihan ulasan positif secara selektif berdasarkan kriteria berikut:

1. Hanya ulasan positif dengan panjang karakter lebih dari 25 karakter yang digunakan.
2. Ulasan ini dianggap lebih informatif karena memiliki kalimat yang lebih panjang.

Setelah seleksi, sebanyak 208 ulasan positif dipilih, sehingga dataset akhir yang digunakan untuk pelatihan dan pengujian terdiri dari 172 ulasan negatif dan 208 ulasan positif. Label sentimen diklasifikasikan sebagai berikut:

1. Positif: Penilaian bintang 4-5 dengan panjang ulasan > 25 karakter.
2. Negatif: Penilaian bintang 1-3.

Langkah ini bertujuan untuk mengatasi ketidakseimbangan data sekaligus memastikan bahwa ulasan yang digunakan memiliki informasi yang cukup untuk analisis sentimen, karena pada penelitian [12] terbukti bahwa Naive Bayes memperoleh nilai performa tidak stabil pada *imbalanced dataset*.

### B. Preprocessing Data

Tahapan preprocessing data dilakukan untuk memastikan bahwa data yang digunakan dalam pelatihan dan pengujian model klasifikasi bersih, seragam, dan siap untuk dianalisis. Proses preprocessing melibatkan beberapa langkah utama berikut:

1. Penghapusan Tanda Baca: Seluruh tanda baca seperti titik, koma, tanda seru, dan lainnya dihapus untuk menyederhanakan teks.
2. Penghapusan Angka: Angka yang tidak relevan dihapus agar teks lebih fokus pada informasi sentimen.
3. Penghapusan Karakter Khusus: Karakter seperti simbol dan emotikon yang tidak memberikan kontribusi signifikan terhadap analisis sentimen dihilangkan.
4. Penghapusan URL: Untuk melakukan antisipasi apabila dalam tulisan terdapat link yang tidak relevan untuk analisis sentimen.
5. Stemming: Kata-kata diubah ke bentuk dasarnya menggunakan algoritma stemming bahasa Indonesia, seperti *Sastrawi*. Misalnya, kata "berjalan" akan diubah menjadi "jalan".
6. *Stopword Removal*: Kata-kata umum yang tidak memiliki makna signifikan dalam analisis

sentimen, seperti "yang", "dan", "atau", dihapus menggunakan daftar stopwords bahasa Indonesia.

7. Tokenisasi: Setiap ulasan dipecah menjadi token atau kata-kata individu untuk mempermudah analisis lebih lanjut.
8. *Lowercasing*: Seluruh teks diubah menjadi huruf kecil untuk memastikan konsistensi, sehingga kata seperti "Rumah" dan "rumah" dianggap sama.

Setelah preprocessing, data teks ulasan diubah menjadi bentuk vektor numerik menggunakan teknik *Term Frequency-Inverse Document Frequency* (TF-IDF). Representasi ini memungkinkan model klasifikasi untuk memahami pola dalam data teks berdasarkan frekuensi kata dan bobotnya dalam dokumen [13]. Tahapan preprocessing ini bertujuan untuk meningkatkan akurasi dan efisiensi model Naive Bayes dalam menganalisis sentimen ulasan rumah sakit.

### C. Pelatihan Metode Klasifikasi

Pada tahap ini, model klasifikasi Naive Bayes dilatih untuk menganalisis sentimen ulasan terhadap Rumah Sakit Muhammadiyah Lamongan. Metode Naive Bayes dipilih karena kemampuannya yang efektif untuk menangani data teks dengan asumsi independensi antar fitur.

Naive Bayes adalah algoritma pembelajaran mesin yang berbasis pada Teorema Bayes, yang digunakan untuk memprediksi kelas suatu data berdasarkan probabilitas bersyarat. Algoritma ini banyak digunakan untuk tugas klasifikasi, seperti analisis sentimen, deteksi spam, dan klasifikasi dokumen. Teorema Bayes digunakan untuk menghitung probabilitas posterior  $P(C|X)$ , yaitu probabilitas suatu data  $X$  termasuk ke dalam kelas  $C$ , berdasarkan probabilitas sebelumnya  $P(C)$  dan  $P(X/C)$ :

$$P(C|X) = \frac{P(X|C) \cdot P(C)}{P(X)} \quad (1)$$

Penjelasan dari Persamaan (1) :

- $P(C|X)$ : Probabilitas posterior, yaitu probabilitas data  $X$  diberikan kelas  $C$ .
- $P(C)$ : Probabilitas prior dari kelas  $C$ .
- $P(X|C)$ : Probabilitas data  $X$  diberikan kelas  $C$ .
- $P(X)$ : Probabilitas data  $X$  secara keseluruhan.

Beberapa varian Naive Bayes digunakan untuk pelatihan [14], yaitu:

1. *Complement Naive Bayes*: Varian ini dirancang untuk menangani ketidakseimbangan kelas pada data. *Complement Naive Bayes* memberikan bobot lebih besar pada fitur yang cenderung muncul di kelas minoritas, sehingga cocok untuk data yang distribusinya tidak merata.
2. *Multinomial Naive Bayes*: Varian ini bekerja dengan baik pada data diskrit seperti frekuensi

kata dalam dokumen. Multinomial Naive Bayes mengasumsikan bahwa fitur merupakan distribusi multinomial [15], sehingga cocok untuk data yang telah diolah menggunakan teknik *Term Frequency-Inverse Document Frequency* (TF-IDF).

3. Bernoulli Naive Bayes: Varian ini menggunakan data biner, di mana setiap fitur menunjukkan ada atau tidaknya suatu kata dalam dokumen [16]. Bernoulli Naive Bayes sering digunakan untuk data teks yang sederhana atau ketika representasi data hanya berupa kehadiran fitur tertentu.

Data yang telah melalui preprocessing dibagi menjadi dua subset, yaitu data pelatihan (70%) dan data pengujian (30%). Pembagian ini dilakukan secara acak untuk memastikan bahwa model dilatih dan diuji pada data yang tidak saling terkait. Setiap varian Naive Bayes dilatih menggunakan data pelatihan. Model mempelajari hubungan antara fitur (TF-IDF) dan label sentimen (positif atau negatif) untuk menghasilkan probabilitas posterior yang digunakan dalam prediksi. Performa masing-masing varian dievaluasi menggunakan data pengujian, dengan mempertimbangkan metrik akurasi, presisi, recall, dan F1-score. Varian dengan performa terbaik akan dipilih sebagai model utama untuk analisis sentimen.

Tahapan ini memastikan bahwa model yang digunakan tidak hanya akurat, tetapi juga andal dalam mengklasifikasikan ulasan masyarakat secara konsisten.

#### D. Pengujian dan Evaluasi

Setelah model klasifikasi Naive Bayes dilatih, tahap selanjutnya adalah pengujian untuk menilai performa model dalam mengklasifikasikan sentimen ulasan. Data pengujian digunakan untuk mengukur kemampuan model dalam mengklasifikasikan ulasan sebagai sentimen positif atau negatif. Evaluasi dilakukan dengan menggunakan beberapa metrik untuk memberikan gambaran yang lebih komprehensif mengenai kinerja model [17], yaitu:

1. Akurasi: Metrik ini mengukur seberapa banyak prediksi model yang benar dibandingkan dengan total prediksi yang dibuat. Akurasi dihitung dengan Persamaan (2) seperti berikut.

$$\text{Akurasi} = \frac{\text{Jumlah Prediksi Benar}}{\text{Jumlah Total Data}} \quad (2)$$

2. Presisi: Mengukur seberapa banyak ulasan yang diklasifikasikan sebagai positif oleh model yang benar-benar positif. Presisi dihitung dengan Persamaan (3)

$$\text{Presisi} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \quad (3)$$

3. Recall: Mengukur seberapa banyak ulasan positif yang sebenarnya benar-benar teridentifikasi oleh model. Recall dihitung dengan Persamaan (4)

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \quad (4)$$

4. F1-Score: Merupakan rata-rata harmonik dari presisi dan recall, yang memberikan nilai yang lebih seimbang antara keduanya. F1-Score dihitung dengan Persamaan (5)

$$\text{F1-Score} = 2 \times \frac{\text{Presisi} \times \text{Recall}}{\text{Presisi} + \text{Recall}} \quad (5)$$

Selain metrik tersebut, Receiver Operating Characteristic (ROC) curve juga digunakan untuk mengevaluasi performa model, terutama dalam konteks klasifikasi biner. ROC adalah grafik yang menggambarkan hubungan antara True Positive Rate (TPR) dan False Positive Rate (FPR) pada berbagai threshold. TPR dihitung sebagai recall, sementara FPR dihitung dengan Persamaan (6).

$$\text{FPR} = \frac{\text{False Positives}}{\text{False Positives} + \text{True Negatives}} \quad (6)$$

Dengan menggunakan ROC curve, kita dapat memvisualisasikan trade-off antara sensitivitas dan spesifisitas model, serta menghitung Area Under the Curve (AUC). Semakin besar nilai AUC (mendekati 1), semakin baik model dalam membedakan antara kelas positif dan negatif.

### III. HASIL DAN PEMBAHASAN

Pada penelitian ini, dilakukan analisis sentimen terhadap ulasan Google Maps Rumah Sakit Muhammadiyah Lamongan menggunakan tiga varian algoritma Naive Bayes: Multinomial Naive Bayes, Bernoulli Naive Bayes, dan Gaussian Naive Bayes. Hasil dari analisis ini disajikan melalui berbagai visualisasi data seperti word cloud, kata-kata umum dalam teks, dan kurva ROC (Receiver Operating Characteristic).

Data ulasan dari Google Maps yang terkumpul melalui scraping berjumlah 380 ulasan dan dikategorikan ke dalam dua kelas sentimen, yaitu positif dan negatif. Untuk memastikan kualitas data, dilakukan tahap pre-processing, yang meliputi penghapusan tanda baca, angka, karakter khusus, URL, serta normalisasi dengan stemming dan stopword removal. Selanjutnya, data teks diubah menjadi representasi numerik menggunakan metode TF-IDF (Term Frequency-Inverse Document Frequency). Representasi TF-IDF digunakan sebagai input untuk pelatihan model Naive Bayes.

#### A. Visualisasi Word Cloud

Untuk memberikan gambaran yang lebih jelas mengenai pola dalam ulasan pasien, digunakan visualisasi word cloud. Word cloud ini menampilkan

kata-kata yang paling sering muncul dalam ulasan pasien berdasarkan sentimen positif dan negatif.

Gambar word cloud ulasan positif seperti pada Gambar 2 menunjukkan bahwa kata-kata seperti "pelayanan," "pasien," "perawat," dan "baik" sering muncul. Kata-kata ini mengindikasikan bahwa aspek pelayanan rumah sakit, terutama yang terkait dengan sikap perawat dan kenyamanan pasien, menjadi sorotan utama dalam ulasan positif. Selain itu, kata-kata seperti "semoga," "terima kasih," dan "Alhamdulillah" mencerminkan apresiasi pasien terhadap pelayanan yang diterima.

Sebaliknya, Gambar word cloud ulasan negatif pada Gambar 3 menunjukkan dominasi kata-kata seperti "tidak," "lama," dan "kurang." Kata-kata ini mengarah pada ketidakpuasan pasien, terutama pada aspek waktu pelayanan dan kualitas pelayanan tertentu. Selain itu, kata-kata seperti "dokter," "antrian," dan "satpam" menunjukkan bahwa keluhan sering terkait dengan pelayanan tenaga medis, antrian yang panjang, dan interaksi dengan staf non-medis.



Gambar 2. Word Cloud untuk Ulasan Positif

Word cloud ini tidak hanya berfungsi sebagai alat visualisasi, tetapi juga menjadi dasar untuk memahami aspek mana yang perlu dipertahankan dan aspek mana yang perlu diperbaiki oleh pihak manajemen rumah sakit.



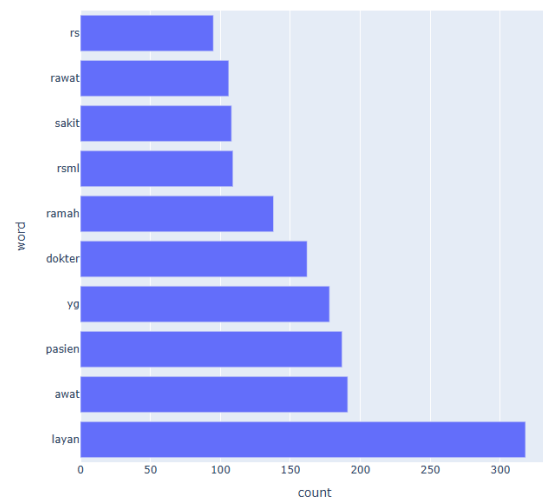
Gambar 3. Word Cloud untuk Ulasan Negatif

#### B. Common Words

Grafik common words seperti pada Gambar 4 memberikan gambaran tentang kata-kata yang paling sering muncul dalam keseluruhan ulasan (positif maupun negatif). Kata yang paling dominan adalah "layan" (pelayanan), diikuti oleh "awat" (merujuk pada "perawatan"), "pasien," dan "dokter".

Kehadiran kata-kata ini mencerminkan fokus utama pasien terhadap aspek pelayanan medis, interaksi dengan dokter, serta pengalaman perawatan secara keseluruhan. Kata-kata seperti "ramah" juga muncul secara signifikan, menunjukkan bahwa keramahan tenaga medis menjadi salah satu perhatian pasien.

Namun, kata "rsml" (Rumah Sakit Muhammadiyah Lamongan) juga sering disebutkan, mengindikasikan bahwa ulasan memiliki fokus spesifik pada institusi ini. Kata seperti "yg" menunjukkan bahasa informal yang digunakan oleh pengulas, memberikan konteks realistis terhadap persepsi pasien dalam format ulasan digital.



Gambar 4. Common Words

Grafik common words ini memberikan beberapa wawasan penting:

1. Fokus pada Pelayanan: Kata-kata seperti "layan," "awat," dan "dokter" menunjukkan bahwa pasien memiliki ekspektasi tinggi terhadap kualitas pelayanan medis, khususnya terkait interaksi langsung dengan dokter dan staf medis.
2. Keramahan Tenaga Medis: Kehadiran kata "ramah" menunjukkan bahwa sikap positif tenaga medis menjadi salah satu aspek yang diapresiasi pasien, meskipun aspek ini juga muncul dalam keluhan pada ulasan negatif.
3. Institusi yang Dikenal: Penyebutan "rsml" secara konsisten mengindikasikan bahwa rumah sakit ini cukup dikenal oleh masyarakat dan menjadi sorotan dalam platform ulasan digital.

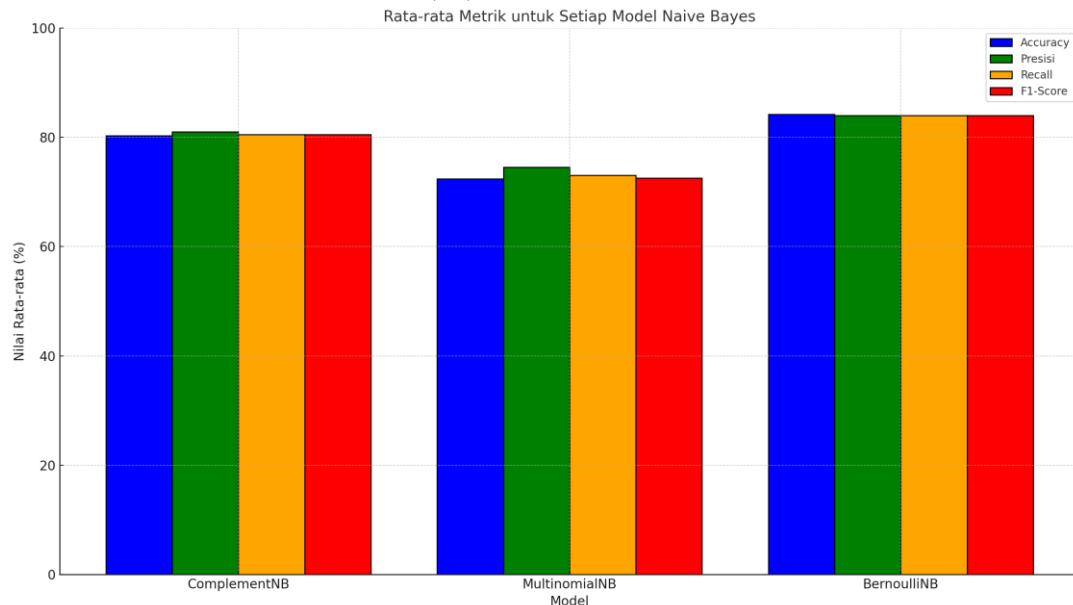
#### C. Implementasi Naïve Bayes

Dari dataset yang sudah dipersiapkan pada tahap Preprocessing Data selanjutnya ditraining menggunakan metode Naïve Bayes. Tahapan implementasi Naive Bayes pada data ini meliputi:

1. Menerapkan TF-IDF

Setelah preprocessing, data teks diubah menjadi representasi numerik menggunakan TF-IDF (Term Frequency-Inverse Document Frequency). TF-IDF menghitung bobot

pentingnya sebuah kata dalam dokumen berdasarkan frekuensi kemunculan kata (TF) dan



Gambar 5. Grafik Hasil Evaluasi Model Naive Bayes

seberapa jarang kata tersebut muncul di semua dokumen (IDF). Contoh sederhana perhitungan TF-IDF seperti pada Persamaan (7) dibawah ini.

$$TF-IDF(w) = TF(w) \times \log \left( \frac{N}{DF(w)} \right) \quad (7)$$

di mana  $w$  adalah kata,  $N$  adalah jumlah total dokumen, dan  $DF(w)$  adalah jumlah dokumen yang mengandung kata  $w$ .

## 2. Split Dataset

Data yang telah direpresentasikan dalam bentuk TF-IDF dibagi menjadi:

- 70% data latih (training data): Digunakan untuk membangun model klasifikasi.
- 30% data uji (testing data): Digunakan untuk menguji performa model.

## 3. Pelatihan Metode Klasifikasi

Data latih digunakan untuk melatih tiga varian model Naive Bayes yaitu Complement Naive Bayes (CNB), Multinomial Naive Bayes (MNB), dan Bernoulli Naive Bayes (BNB). Proses ini melibatkan langkah-langkah sebagai berikut:

- Inisialisasi Model: Ketiga model Naive Bayes (ComplementNB, MultinomialNB, BernoulliNB) diinisialisasi dengan parameter default atau menggunakan faktor smoothing seperti  $\alpha=1$ .
- Input Data: Data yang telah diproses dalam bentuk TF-IDF matrix dimasukkan sebagai input untuk model.
- Probabilitas Kelas: Model menghitung probabilitas prior untuk setiap kelas (positif dan negatif).
- Pelatihan Model: Model mempelajari distribusi probabilitas kata (likelihood) dalam setiap kelas berdasarkan data latih.

Perhitungan probabilitas dalam Naive Bayes menggunakan rumus seperti Persamaan (8).

$$P(C|X) = P(C) \prod_{i=1}^n P(w_i|C) \quad (8)$$

di mana:

- $P(C)$ : Probabilitas awal kelas (prior).
- $P(w_i|C)$ : Probabilitas setiap kata  $w_i$  muncul pada kelas  $C$ .
- $P(C|X)$ : Probabilitas kelas  $C$  diberikan dokumen  $X$ .

Probabilitas untuk setiap kelas dihitung, dan dokumen diklasifikasikan ke dalam kelas dengan nilai probabilitas tertinggi.

## D. Hasil Evaluasi Model Naive Bayes

Hasil pengujian seperti pada Tabel 1 dan Gambar 5 menunjukkan bahwa model BernoulliNB memiliki performa terbaik dengan akurasi tertinggi sebesar 84,21%. Model ini juga menunjukkan presisi dan recall yang konsisten untuk kedua kelas (negatif dan positif), yaitu sekitar 84%. F1-Score rata-rata untuk BernoulliNB mencapai 84%, yang mencerminkan keseimbangan antara kemampuan model dalam mendeteksi ulasan positif maupun negatif. Di sisi lain, ComplementNB memberikan performa yang cukup kompetitif dengan akurasi sebesar 80,26%. Model ini unggul dalam mendeteksi ulasan positif dengan recall yang lebih besar menjadikannya alternatif yang baik jika aspek ini menjadi prioritas. Namun, model MultinomialNB memiliki performa yang paling lemah, dengan akurasi hanya sebesar 72,37%. Presisi dan recall untuk kelas negatif pada model ini jauh lebih

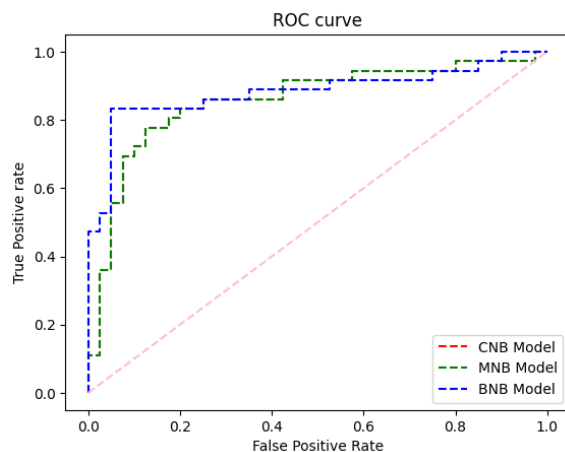
rendah dibandingkan dengan model lainnya, sehingga menghasilkan F1-Score yang kurang memuaskan, terutama untuk kelas negatif. Secara keseluruhan, BernoulliNB menjadi pilihan terbaik untuk analisis sentimen ini, diikuti oleh ComplementNB sebagai alternatif yang cukup andal, sementara MultinomialNB memerlukan peningkatan lebih lanjut dalam hal penanganan ulasan negatif.

Dapat diamati pada Gambar 6 model BernoulliNB memiliki kurva ROC yang paling dominan dengan nilai AUC (Area Under the Curve) yang lebih besar dibandingkan model lainnya, menunjukkan kemampuan terbaik dalam memisahkan kelas positif dan negatif dengan tingkat kesalahan yang lebih kecil. ComplementNB berada di posisi kedua, dengan kurva yang mendekati BernoulliNB, namun sedikit lebih rendah, mencerminkan performa yang cukup baik tetapi tidak seoptimal BernoulliNB. Sementara itu, MultinomialNB menunjukkan kurva ROC yang lebih rendah, mengindikasikan bahwa model ini memiliki tingkat kesalahan yang lebih tinggi dan kurang mampu menangani trade-off antara sensitivitas dan spesifisitas pada berbagai nilai threshold.

Tabel 2. Media Bibit

Model	Akurasi (%)	Presisi (%)	Recall (%)	F1-Score (%)
ComplementNB	80.26	81	80.5	80.5
MultinomialNB	72.37	74.5	73	72.5
BernoulliNB	84.21	84	84	84

Dari segi True Positive Rate (TPR) dan False Positive Rate (FPR), model BernoulliNB menunjukkan TPR yang lebih tinggi pada sebagian besar nilai FPR, yang berarti model ini lebih efektif dalam menangkap ulasan positif sambil menjaga tingkat kesalahan prediksi ulasan negatif tetap rendah. ComplementNB juga memiliki kinerja yang cukup stabil, meskipun tidak setinggi BernoulliNB. Sebaliknya, MultinomialNB menunjukkan peningkatan TPR yang kurang stabil dan FPR yang lebih tinggi, menandakan kelemahan model dalam memisahkan kelas dengan baik.



Gambar 6. ROC Curve

Secara keseluruhan, BernoulliNB adalah model dengan performa terbaik berdasarkan kurva ROC, karena memberikan keseimbangan optimal antara sensitivitas dan spesifisitas. ComplementNB dapat dijadikan alternatif yang layak dengan performa mendekati BernoulliNB, sementara MultinomialNB memerlukan optimasi lebih lanjut, seperti tuning parameter atau modifikasi preprocessing data, agar dapat meningkatkan kinerjanya. Hasil ini mendukung bahwa BernoulliNB sangat cocok untuk aplikasi analisis sentimen, terutama dalam skenario di mana penting untuk meminimalkan kesalahan prediksi ulasan negatif sebagai positif atau sebaliknya.

#### IV. KESIMPULAN

Penelitian ini bertujuan untuk membandingkan performa tiga model Naive Bayes, yaitu ComplementNB, MultinomialNB, dan BernoulliNB, dalam analisis sentimen terhadap ulasan publik. Berdasarkan hasil evaluasi, BernoulliNB menunjukkan performa terbaik dengan akurasi tertinggi sebesar 84,21% serta nilai presisi, recall, dan F1-Score yang konsisten tinggi. ComplementNB menjadi alternatif yang kompetitif dengan akurasi 80,26%, meskipun sedikit lebih rendah dibandingkan BernoulliNB. Di sisi lain, MultinomialNB menunjukkan performa paling lemah dengan akurasi 72,37%, mengindikasikan bahwa model ini kurang efektif untuk dataset analisis sentimen tanpa adanya optimasi lebih lanjut. Selain itu, analisis menggunakan kurva ROC menunjukkan bahwa BernoulliNB memiliki AUC tertinggi, memperkuat kesimpulan bahwa model ini memberikan keseimbangan optimal antara sensitivitas dan spesifisitas.

Terdapat beberapa saran yang bisa digunakan untuk penelitian selanjutnya:

1. Optimasi Dataset: Penelitian selanjutnya disarankan untuk menggunakan teknik preprocessing yang lebih kompleks, seperti stemming atau lemmatization, guna meningkatkan kualitas fitur teks yang digunakan dalam model.
2. Eksplorasi Model Lain: Penelitian lebih lanjut dapat mempertimbangkan model lain seperti Support Vector Machine (SVM) atau model berbasis deep learning seperti LSTM untuk meningkatkan akurasi analisis sentimen.
3. Evaluasi pada Dataset Berbeda: Perlu dilakukan pengujian pada dataset dengan karakteristik yang berbeda, seperti dataset dengan distribusi sentimen yang tidak seimbang, untuk mengevaluasi generalisasi model yang digunakan.

## DAFTAR PUSTAKA

- [1] J. D. C. Aruan, B. Rahayudi, and A. Ridok, "Analisis Sentimen Opini Masyarakat terhadap Pelayanan Rumah Sakit Umum Daerah menggunakan Metode Support Vector Machine dan Term Frequency – Inverse Document Frequency," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 6, no. 5, pp. 2072–2078, 2022.
- [2] 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. 11, no. 5, pp. 993–1000, Oct. 2024, doi: 10.25126/jtiik.1077979.
- [3] A. C. T. Angel, V. H. Pranatawijaya, and W. Widiatry, "Analisis Sentimen dan Emosi dari Ulasan Google Maps Untuk Layanan Rumah Sakit di Palangka Raya Menggunakan Machine Learning," *KONSTELASI Konvergensi Teknol. dan Sist. Inf.*, vol. 4, no. 1, pp. 35–49, 2024, doi: 10.24002/konstelasi.v4i1.8924.
- [4] C. Prakoso and A. Hermawan, "Perbandingan Model Machine Learning dalam Analisis Sentimen Ulasan Pengunjung Keraton Yogyakarta pada Google Maps," *Kaji. Ilm. Inform. dan Komput.*, vol. 4, no. 3, pp. 1292–1302, 2023, doi: 10.30865/KLIK.V4I3.1419.
- [5] A. P. Giovani, A. Ardiansyah, T. Haryanti, L. Kurniawati, and W. Gata, "Analisis Sentimen Aplikasi Ruang Guru Di Twitter Menggunakan Algoritma Klasifikasi," *J. Teknoinfo*, vol. 14, no. 2, p. 115, Jul. 2020, doi: 10.33365/jti.v14i2.679.
- [6] S. Samsir, A. Ambiyar, U. Verawardina, F. Edi, and R. Watrionthos, "Analisis Sentimen Pembelajaran Daring Pada Twitter di Masa Pandemi COVID-19 Menggunakan Metode Naïve Bayes," *J. MEDIA Inform. BUDIDARMA*, vol. 5, no. 1, pp. 157–163, Jan. 2021, doi: 10.30865/MIB.V5I1.2580.
- [7] Syahril Dwi Prasetyo, Shofa Shofiah Hilabi, and Fitri Nurapriani, "Analisis Sentimen Relokasi Ibukota Nusantara Menggunakan Algoritma Naïve Bayes dan KNN," *J. KomtekInfo*, vol. 10, pp. 1–7, 2023, doi: 10.35134/komtekinfo.v10i1.330.
- [8] B. Gunawan, H. S. Pratiwi, and E. E. Pratama, "Sistem Analisis Sentimen pada Ulasan Produk Menggunakan Metode Naive Bayes," *J. Edukasi dan Penelit. Inform.*, vol. 4, no. 2, p. 113, Dec. 2018, doi: 10.26418/jp.v4i2.27526.
- [9] Merinda Lestandy, Abdurrahim Abdurrahim, and Lailis Syafa'ah, "Analisis Sentimen Tweet Vaksin COVID-19 Menggunakan Recurrent Neural Network dan Naïve Bayes," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 4, pp. 802–808, 2021, doi: 10.29207/resti.v5i4.3308.
- [10] Y. Liu and Y. F. B. Wu, "Early detection of fake news on social media through propagation path classification with recurrent and convolutional networks," *32nd AAAI Conf. Artif. Intell. AAAI 2018*, pp. 354–361, 2018.
- [11] Omkarcloud, "GitHub - omkarcloud/botasaurus: The All in One Framework to build Awesome Scrapers." <https://github.com/omkarcloud/botasaurus> (accessed Dec. 10, 2024).
- [12] Ericha Apriliyani and Y. Salim, "Analisis performa metode klasifikasi Naïve Bayes Classifier pada Unbalanced Dataset," *Indones. J. Data Sci.*, vol. 3, no. 2, pp. 47–54, 2022, doi: 10.56705/ijodas.v3i2.45.
- [13] K. Tri Putra, M. Amin Hariyadi, and C. Crysdian, "Perbandingan Feature Extraction Tf-Idf Dan Bow Untuk Analisis Sentimen Berbasis Svm," *J. Cahaya MAndalika*, p. 1449, 2023.
- [14] S. Xu, "Bayesian Naïve Bayes classifiers to text classification," *J. Inf. Sci.*, vol. 44, no. 1, pp. 48–59, 2018, doi: 10.1177/0165551516677946.
- [15] P. R. Alvita Wagiswari D, I. Susilawati, and A. Witanti, "Analisis Sentimen pada Komentar Aplikasi MyPertamina dengan Metode Multinomial Naive Bayes," *ForAI J. Informatics Artif. Intell. J.*, vol. 1, no. 1, pp. 10–19, 2023, [Online]. Available: <https://jurnal.forai.or.id/index.php/forai/article/view/4>.
- [16] A. Yodi Prayoga, A. Id Hadiana, and F. Rakhmat Umbara, "Deteksi Hoax pada Berita Online Bahasa Inggris Menggunakan Bernoulli Naïve Bayes dengan Ekstraksi Fitur Tf-Idf," *J. Syntax Admiration*, vol. 2, no. 10, pp. 1808–1823, 2021, doi: 10.46799/jsa.v2i10.327.
- [17] A. Zheng, *Evaluating Machine Learning Algorithms: A Beginner's Guide to Key Concepts and Pitfalls*. 2015.