

Sentiment Analysis customer Towards Cinema Services in Semarang Using Naive Bayes Classifier on Google Reviews

Husni Brian Maulana ¹, Maya Rini Handayani ², Khothibul Umam ^{3*}

Teknologi Informasi, Fakultas Sains dan Teknologi, UIN Walisongo Semarang, Indonesia

2208096091@student.walisongo.ac.id¹, maya@walisongo.ac.id², khothibul_umam@walisongo.ac.id³

Article Info

Article history:

Received 2025-08-29

Revised 2025-09-25

Accepted 2025-10-21

Keyword:

Sentiment, Naive Bayes Classifier, Google Reviews, Cinema, Semarang.

ABSTRACT

The development of the entertainment industry, especially in the field of cinema, encourages every service provider to continuously maintain the quality of their services. One method of assessing customer satisfaction is through sentiment testing. The main objective of this study is to examine customer sentiment towards cinema services in Semarang by applying the Naive Bayes Classifier method. The research data was taken from 600 customer reviews on Google Review, which were then divided into two groups: training data consisting of 480 reviews (80%) and testing data consisting of 120 reviews (20%). Before the classification process, the data underwent pre-processing stages involving data cleaning, case folding, tokenization, stopword removal, and stemming, followed by data labeling into two sentiment categories, namely positive and negative. This study took five cinemas as objects, namely CitraXXI, Cinépolis Java Mall, Paragon XXI, XXI Uptown Mall, and XXI DP Mall. The classification results show that the Naive Bayes algorithm is able to group sentiments quite well, with model accuracy ranging from 0.90 to 0.94. Of the five cinemas, Cinépolis Java Mall achieved the highest accuracy, which was 0.94.



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

I. PENDAHULUAN

Sentiment analysis, atau yang sering disebut juga sebagai penggalian opini, merupakan sebuah studi Kajian yang dimana mengevaluasi pendapat, pemikiran, dan persepsi seseorang terhadap berbagai tema, objek, produk, atau layanan [1]. "Teknik ini banyak digunakan di berbagai bidang, seperti pemasaran, layanan pelanggan, dan media sosial, untuk memahami persepsi masyarakat serta mendukung pengambilan keputusan yang berbasis data.

Google Reviews dipilih sebagai sumber data karena salah satu platform ulasan yang paling kredibel dan banyak digunakan masyarakat. Setiap pengguna yang memberikan ulasan harus menggunakan akun Google aktif, sehingga meminimalisasi manipulasi data. Selain itu, Google Reviews secara luas diakses oleh masyarakat dari berbagai kalangan, analisis sentimen berbasis google review juga mampu menggambarkan persepsi pelanggan secara real-time. Ulasan pelanggan di platform digital dapat menjadi sumber informasi berharga mengenai tingkat kepuasan, keluhan, maupun keunggulan layanan

Di dunia bisnis yang semakin kompetitif ini, tingkat persaingan sangatlah tinggi. Masing-masing pesaing menyediakan beragam produk dan layanan, dengan persaingan yang mencakup faktor harga, mutu, promosi, serta berbagai upaya untuk memenuhi tingkat kepuasan pelanggan yang berbeda di tiap perusahaan [2]. Industri hiburan, termasuk bioskop, mengalami tekanan serupa. Bioskop tidak hanya bersaing dalam hal kualitas tayangan, tetapi juga dalam pengalaman menonton yang ditawarkan kepada penonton. Di Indonesia, bioskop termasuk ke dalam subsektor hiburan yang telah mengalami perkembangan sejak lama. [3]. Sebagai salah satu bagian dari industri hiburan, bioskop memiliki peran penting dalam menyediakan hiburan dan pengalaman sosial bagi masyarakat.

Hasil penelitian ini tidak hanya memberikan kontribusi akademis berupa evaluasi performa metode klasifikasi teks, tetapi juga menawarkan manfaat praktis bagi pengelola bioskop. Informasi mengenai kecenderungan sentimen pelanggan dapat dijadikan dasar dalam menyusun strategi pemasaran, meningkatkan kualitas layanan, serta memperbaiki aspek yang sering dikeluhkan.

Sebagai pusat pemerintahan Jawa Tengah, Kota Semarang menunjukkan perkembangan bertahap dalam penyediaan fasilitas bioskop. Meskipun bukan yang paling pesat Di antara kota-kota besar di Indonesia, Semarang juga menunjukkan perkembangan dalam dunia perfilman, terbukti dengan kepercayaannya menjadi tuan rumah Festival Film Indonesia pada tahun 2013. [4].

Dalam proses klasifikasi ulasan, tersedia berbagai metode, dengan Naïve Bayes sebagai salah satu pilihan yang sering diterapkan, Naïve Bayes merupakan teknik klasifikasi yang menggunakan pendekatan probabilistik sederhana, berlandaskan pada Teorema Bayes dan menganggap tiap fitur tidak saling bergantung [5]. Naïve Bayes terbukti efektif dalam berbagai aplikasi, khususnya dalam analisis teks. Metode ini dapat digunakan untuk mengklasifikasikan ulasan berdasarkan polaritas sentimennya, baik positif maupun negatif.

Tujuan dari penelitian ini adalah mengklasifikasikan ulasan pelanggan bioskop di Kota Semarang yang diperoleh melalui Google Reviews ke dalam dua kategori sentimen, yaitu *positif* dan *negatif*, serta membandingkan hasil analisis antar-bioskop untuk mengetahui kecenderungan ulasan pada masing-masing bioskop. Dengan demikian, penelitian ini tidak hanya memberikan gambaran mengenai persepsi dominan pelanggan terhadap layanan bioskop, tetapi juga dapat menjadi dasar bagi pihak manajemen dalam melakukan evaluasi dan peningkatan kualitas layanan secara lebih tepat sasaran.

Penelitian ini tidak hanya berfokus pada penggalian sentimen pelanggan terhadap layanan bioskop di Kota Semarang, tetapi juga untuk mengevaluasi performa metode Naïve Bayes dalam mengklasifikasikan teks ulasan. Evaluasi ini penting dilakukan karena setiap domain data memiliki karakteristik berbeda, sehingga kinerja algoritma perlu diuji secara spesifik agar hasil klasifikasi dapat diandalkan.

Penelitian yang dilakukan oleh Ahmad Afif [6] dengan menganalisis Keunikan Hotel menunjukkan hasil evaluasi menggunakan Naïve Bayes dengan akurasi 0,75 % berdasarkan 120 data, dengan rasio data training dan testing 80:20.

Penelitian yang dilakukan oleh Margaretha Inya Tuku, Elfira Umar dan Alexander Adis [7] dengan menganalisis objek Ella Hotel Tambolaka menunjukkan hasil evaluasi menggunakan Naïve Bayes dengan akurasi 90,71 % berdasarkan 128 data, dengan rasio data training dan testing 80:20. Penelitian yang dilakukan oleh Adinda Cahya Kamilla, Natasya Priyani, dan Viktor Handrianus Pranatawijaya [8] dengan menganalisis objek film agak laen menunjukkan hasil evaluasi menggunakan Naïve Bayes dengan akurasi 78 % berdasarkan 195 data, dengan rasio data training dan testing 80:20. Penelitian yang dilakukan oleh Fithri Selva Jumeilah [9] dengan menganalisis objek Jasa Ekspedisi JNE menunjukkan hasil evaluasi menggunakan Naïve Bayes dengan akurasi 85 % berdasarkan 1725 data, dengan rasio data training dan testing 70:30.

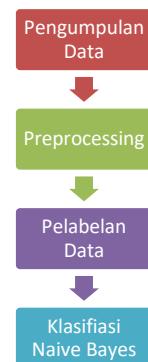
Penelitian dari Akhmad Pandhu Wijaya dan Heru Agus Santoso [10] menganalisis objek Konten E-Government menunjukkan hasil evaluasi menggunakan Naïve Bayes dengan klasifikasi akurasi 85 % berdasarkan 260 data, dengan rasio data training dan testing 80:20. Penelitian yang dilakukan oleh Nurdin*, M. Suhendri, Yesy Afrilia, Rizal [11] dengan menganalisis objek Karya Ilmiah menunjukkan hasil evaluasi menggunakan Naïve Bayes dengan rata rata akurasi 86,68 % berdasarkan 170 data, dengan data training 150 data dan testing 20 data. Penelitian yang dilakukan oleh Sifa Melina Salsabila1, Aang Alim Murtopo,Nurul Fadhilah [12] dengan menganalisis objek Pelanggan Tokopedia menunjukkan hasil evaluasi menggunakan Naïve Bayes dengan klasifikasi akurasi 95,10% berdasarkan 714 data, dengan rasio data training dan testing 80:20.

Sejumlah penelitian sebelumnya mengenai analisis sentimen pada sektor layanan hotel, e-commerce, maupun jasa ekspedisi Namun minimnya kajian serupa pada sektor hiburan khususnya bioskop di Kota Semarang. Padahal, bioskop memiliki peran strategis tidak hanya sebagai penyedia hiburan, tetapi juga sebagai sektor bisnis yang sangat dipengaruhi oleh pengalaman pelanggan. Minimnya penelitian di ranah ini menimbulkan kesenjangan riset yang penting untuk diisi, karena ulasan pelanggan di platform digital seperti Google Reviews dapat memberikan gambaran nyata mengenai kualitas layanan bioskop.

II. METODE

A. Pengumpulan data

Penelitian ini memanfaatkan data komentar yang diambil dari Google Maps, dengan total 600 komentar yang diperoleh melalui teknik scraping. Data tersebut kemudian Terbagi ke dalam dua klasifikasi, yakni training data dan testing data. Dengan metode scraping ini, informasi yang diperoleh mampu menggambarkan kebutuhan masyarakat secara langsung dan real-time.



Gambar 1. Alur Penelitian

B. Preprocessing

Tujuan dari tahap ini adalah untuk membuat data teks menjadi lebih bersih, lebih terstruktur, dan lebih siap digunakan untuk analisis lanjutan. Tahap preprocessing

mencakup proses pengumpulan data. scraping Google Maps, Case folding, Cleaning, Tokenisasi, Normalization, Stop Removal, Stemming dan labeling

- 1) *Case Folding*: Case folding merupakan tahap dalam pemrosesan teks yang digunakan untuk menyeragamkan format huruf [13]. Dengan menyeragamkan semua huruf pada data teks ke bentuk lowercase..
- 2) *Cleaning*: Cleaning data adalah proses membersihkan dataset dengan menghapus kolom yang tidak diperlukan, menghilangkan karakter-karakter khusus, serta mengatasi data yang duplikat [14].
- 3) *Tokenisasi*: Tokenizing adalah proses memecah teks menjadi unit-unit kecil seperti kata [15].
- 4) *Normalization*: di tahap ini, komentar yang memiliki kata bermakna negatif akan mempengaruhi perubahan skor sentimennya. . Contohnya kata-kata negasi tersebut antara lain “bukan”, “tidak”, “enggak”, “ga”, “jangan”, “nggak”, “tak”, dan “gak” [16].
- 5) *Stop Removal*: Stopword removal merupakan tahap dalam pemrosesan teks yang bertujuan untuk menghilangkan kata-kata yang tidak memberikan kontribusi besar terhadap arti kalimat [17].
- 6) *Stemming*: Stemming merupakan teknik untuk mengkonversi kata turunan ke bentuk asalnya, yaitu kata dasar, dengan memperhatikan aturan-aturan khusus [18].

C. Pelabelan data

Dalam pelabelan data, ulasan diklasifikasikan ke dalam salah satu dari dua kelompok sentimen, yaitu sentimen positif atau negatif [19] . Proses pelabelan pada penelitian ini dilakukan secara otomatis,

D. Pengujian Metode Naïve Bayes Classifier

Setelah melewati tahapan preprocessing dan pelabelan data, proses pengujian dilanjutkan dengan menerapkan metode Naïve Bayes Classifier. Pengujian ini bertujuan untuk mengevaluasi seberapa akurat model dalam mengklasifikasikan komentar Google Maps ke dalam salah satu dari dua kelompok sentimen, yaitu sentimen positif atau negatif. Banyak penelitian sebelumnya (misalnya pada hotel, e-commerce, jasa ekspedisi) juga menggunakan Naïve Bayes sebagai metode utama, sehingga penggunaan algoritma ini memungkinkan perbandingan hasil dengan studi-studi terdahulu.

Agar menganalisis yang dilakukan menghasilkan nilai yang akurat, penting untuk membagi data yang dikategorikan ke dalam dua utama: Training Data sebanyak 80% dan Testing Data sebanyak 20%. Hasil Pembagian ini sangat krusial untuk memastikan model yang dibangun dapat diuji dan dievaluasi secara objektif, juga memiliki potensi untuk mengenali pola pada data yang sebelumnya tidak dikenal atau belum diproses. Evaluasi terhadap performa klasifikasi

dilakukan dengan menghitung nilai seperti akurasi, presisi, recall, dan f-measure [20].

$$\text{Accuracy} = \frac{(TP+TN)}{(TP+TN+FP+FN)}$$

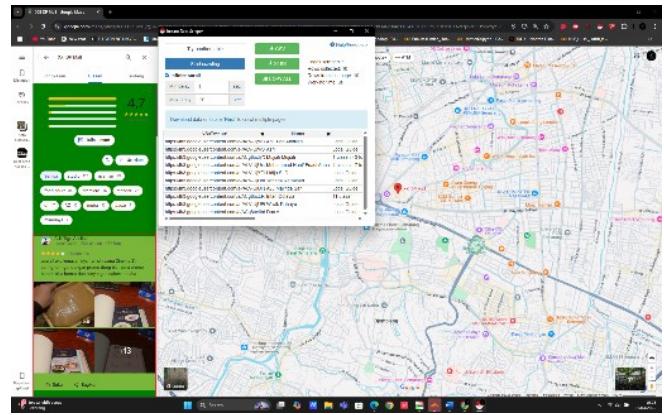
$$\text{Presisi} = \frac{TP}{(TP+FP)}$$

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

$$\text{f-measure} = 2 \frac{\text{Presisi} \times \text{Recall}}{\text{Presisi} + \text{Recall}}$$

III. HASIL DAN PEMBAHASAN

A. Pengumpulan data



gambar 2. Proses Scraping Data

Data yang dikumpulkan berupa komentar pada rating di aplikasi Google Maps tanpa memiliki label nilai positif atau negatif. Hasil pengumpulan data berjumlah 600 komentar yang diperoleh melalui proses scraping otomatis menggunakan ekstensi Chrome Instant Web Scraper.

TABEL 1
SAMPLE DATA

No	Nickname	Comment
1.	Ilham Rashif Sasongko	Pelayanan di XXI Paragon oke banget. Aku sempat ketinggalan barang pas nonton, tapi stafnya cepet tanggap dan bantuin sampai ketemu. Nggak ribet, semua dilayani dengan santai dan profesional. Tempatnya juga nyaman, selalu jadi pilihan buat nonton. Thank you, tim XXI!
2.	Maulim Tobing	Gede sih, tapi tidak sekomplit Citraland simpang 5
3.	Velizca mackenzie	Biskop terpanas 1 semarang, kualitas tempat kayak udh ga dipedulii lagi, nonton film 2 jam full kepanasan sampe liat beberapa org pulang 🤦‍♂️ memang ga ada SOP untuk AC ya ?

4.	Alya Aulia Nurdin	salah satu bioskop favoritku di semarang, ga terlalu rame, super comfy. cuman sayangnya terakhir pas ke sini nonton captain america yang imax tetep kedengeran suara tukang lagi renovasi, beberapa dinding juga ada yang basah bekas bocor
5.	Kristiana S	Watching Jumbo with my sister and her four children. As always, a happy experience every time I visited this place. For yesterday, the place was not crowded. The movie was great. One of the best movies I've watched this year.
...
600	Dida Dianar	Lebih enak ini daripada yang sebelah hehehe

Tabel 1 merupakan sampel data yang diperoleh dari hasil scraping menggunakan ekstensi Chrome Instant Web Scraper dari Bioskop Paragon XXI, selanjutnya akan diolah pada tahap pre-processing.

B. Preprocessing

Tahapan Pertama sebelum klasifikasi yaitu Preprocessing. Adapun beberapa tahap dalam Cleaning dan Preprocessing, antara lain Case Folding, Tokenization, Normalization, Stopword Removal dan Stemming.

TABEL 2
PREPROCESSING

No	Comment	Case Folding
1.	Pelayanan di XXI Paragon oke banget. Aku sempat ketinggalan barang pas nonton, tapi stafnya cepet tanggap dan bantuin sampai ketemu. Nggak ribet, semua dilayani dengan santai dan profesional. Tempatnya juga nyaman, selalu jadi pilihan buat nonton. Thank you, tim XXI!	pelayanan di xxi paragon oke banget aku sempat ketinggalan barang pas nonton, tapi stafnya cepet tanggap dan bantuin sampai ketemu. Nggak ribet, semua dilayani dengan santai dan profesional. Tempatnya juga nyaman, selalu jadi pilihan buat nonton. Thank you, tim XXI!

Tabel 2 merupakan hasil proses Case Folding yang bertujuan untuk menyeragamkan format huruf pada teks ulasan. Pada tahap ini, seluruh huruf diubah menjadi huruf kecil (lowercase) sehingga tidak ada perbedaan antara huruf kapital dan huruf kecil. Proses ini penting untuk menghindari kesalahan pengolahan data akibat perbedaan bentuk huruf, sehingga teks menjadi lebih konsisten dan siap digunakan pada tahap pembersihan (cleaning) berikutnya.

TABEL 3
CLEANING DATA

No	Comment	Clean Text
1.	Pelayanan di XXI Paragon oke banget. Aku sempat ketinggalan barang pas nonton, tapi stafnya cepet tanggap dan bantuin sampai ketemu. Nggak ribet, semua dilayani dengan santai dan profesional. Tempatnya juga nyaman, selalu jadi pilihan buat nonton. Thank you, tim XXI!	pelayanan di xxi paragon oke banget aku sempat ketinggalan barang pas nonton tapi stafnya cepet tanggap dan bantuin sampai ketemu. Nggak ribet semua dilayani dengan santai dan profesional. Tempatnya juga nyaman selalu jadi pilihan buat nonton thank you tim xxi

Tabel 3 merupakan hasil proses Cleaning meliputi menhilangkan tanda baca, simbol-simbol, emoji dan angka. Selanjutnya, seluruh karakter teks diubah menjadi huruf kecil, agar data tersusun dengan baik, terorganisir, dan dapat digunakan dalam tahap analisis selanjutnya.

TABEL 4
TOKENIZING

No	Comment	Tokenizing
1.	Pelayanan di XXI Paragon oke banget. Aku sempat ketinggalan barang pas nonton, tapi stafnya cepet tanggap dan bantuin sampai ketemu. Nggak ribet, semua dilayani dengan santai dan profesional. Tempatnya juga nyaman, selalu jadi pilihan buat nonton. Thank you, tim XXI!	['pelayanan', 'di', 'xxi', 'paragon', 'oke', 'banget', 'aku', 'sempat', 'ketinggalan', 'barang', 'pas', 'nonton', 'tapi', 'stafnya', 'cepet', 'tanggap', 'dan', 'bantuin', 'sampai', 'ketemu', 'nggak', 'ribet', 'semua', 'dilayani', 'dengan', 'santai', 'dan', 'profesional', 'tempatnya', 'juga', 'nyaman', 'selalu', 'jadi', 'pilihan', 'buat', 'nonton', 'thank', 'you', 'tim', 'xxi']

Selanjutnya Preprocessing yaitu Tokenizing, pada Tabel 4 dapat diketahui proses tersebut yaitu memecahkan kalimat sebagai sebuah token agar bisa di proses ke tahap selanjutnya.

TABEL 5
NORMALIZATION

No	Comment	Normalization
1.	Pelayanan di XXI Paragon oke banget. Aku sempat ketinggalan barang pas nonton, tapi stafnya cepet tanggap dan bantuin sampai ketemu.	['pelayanan', 'di', 'xxi', 'paragon', 'oke', 'banget', 'aku', 'sempat', 'ketinggalan', 'barang', 'pas', 'nonton', 'tapi', 'stafnya', 'cepet', 'tanggap', 'dan', 'bantuin', 'sampai', 'ketemu', 'nggak', 'ribet', 'semua', 'dilayani', 'dengan', 'santai', 'dan', 'profesional', 'tempatnya', 'juga', 'nyaman', 'selalu', 'jadi', 'pilihan', 'buat', 'nonton', 'thank', 'you', 'tim', 'xxi']

Nggak ribet, semua dilayani dengan santai dan profesional. Tempatnya juga nyaman, selalu jadi pilihan buat nonton. Thank you, tim XXI!	'dengan', 'santai', 'dan', 'profesional', 'tempatnya', 'juga', 'nyaman', 'selalu', 'jadi', 'pilihan', 'buat', 'nonton', 'thank', 'you', 'tim', 'xxi']
--	---

Tabel 5 yakni proses Normalization Setiap komentar yang mengandung kata-kata dengan makna negasi akan mengalami perubahan pada nilai sentimennya. Pada proses ini, Convert negation diterapkan ketika sebuah kata positif didahului oleh kata negasi, sehingga maknanya berubah menjadi negatif, begitu juga jika kata negatif didahului oleh negasi akan menjadi positif.

TABEL 6
STOPWARD REMOVAL

No	Comment	Stopward Removal
1.	Pelayanan di XXI Paragon oke banget. Aku sempat ketinggalan barang pas nonton, tapi stafnya cepet tanggap dan bantuin sampai ketemu. Nggak ribet, semua dilayani dengan santai dan profesional. Tempatnya juga nyaman, selalu jadi pilihan buat nonton. Thank you, tim XXI!	['pelayanan', 'xxi', 'paragon', 'oke', 'banget', 'aku', 'sempat', 'ketinggalan', 'barang', 'pas', 'nonton', 'stafnya', 'cepet', 'tanggap', 'bantuin', 'ketemu', 'ribet', 'semua', 'dilayani', 'santai', 'profesional', 'tempatnya', 'nyaman', 'selalu', 'jadi', 'pilihan', 'buat', 'nonton', 'thank', 'you', 'tim', 'xxi']

Pada Tabel 6 adalah proses Stopward Removal, Tahapan ini difokuskan pada pengambilan kata-kata penting dalam komentar tersebut dan Membuang kata-kata yang tidak penting atau tidak memberikan kontribusi makna.

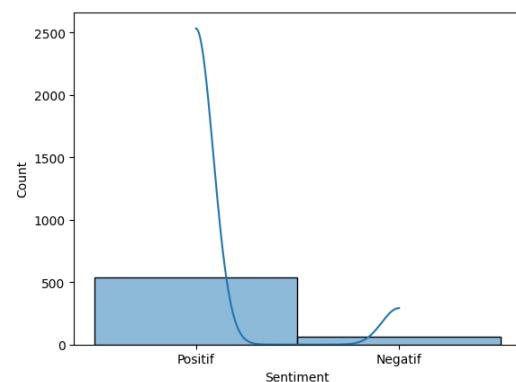
TABEL 7
STEMMING

No	Comment	Stemmed
1.	Pelayanan di XXI Paragon oke banget. Aku sempat ketinggalan barang pas nonton, tapi stafnya cepet tanggap dan bantuin sampai ketemu. Nggak ribet, semua dilayani dengan santai dan profesional. Tempatnya juga nyaman, selalu jadi pilihan buat nonton. Thank you, tim XXI!	['layan', 'xxi', 'paragon', 'oke', 'banget', 'aku', 'sempat', 'tinggal', 'barang', 'pas', 'nonton', 'staf', 'cepet', 'tanggap', 'bantuin', 'ketemu', 'ribet', 'semua', 'layan', 'santai', 'profesional', 'tempat', 'nyaman', 'selalu', 'jadi', 'pilih', 'buat', 'nonton', 'thank', 'you', 'tim', 'xxi']]

Selanjutnya Tabel 7 Proses Stemming bertujuan untuk Menyederhanakan kata ke bentuk asalnya dengan menghilangkan bagian akhir kata menjadi bentuk dasar atau akar kata dengan menghapus akhiran yang tidak relevan, seperti akhiran jamak, atau bentuk lainnya

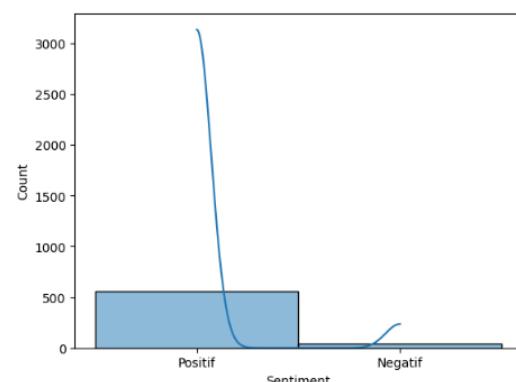
C. Pelabelan data

Pada tahap ini, setiap komentar yang diperoleh dari Google Maps diberi label sesuai dengan makna atau nuansa sentimen yang dikandungnya, yaitu positif atau negatif. Proses pelabelan dapat dilakukan secara manual maupun otomatis. Dalam penelitian ini, pelabelan dilakukan secara manual untuk memastikan ketepatan dan konsistensi dalam penentuan kategori sentimen. Berikut adalah hasil pelabelan untuk lima bioskop di semarang.



Gambar 3. Hasil pelabelan Bioskop CITRAXXI

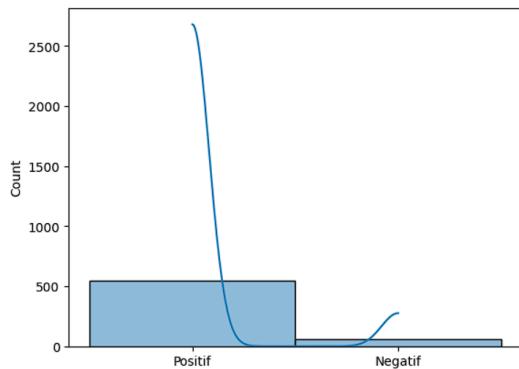
Gambar 3 memperlihatkan hasil pelabelan komentar dari Bioskop CITRAXXI setelah melalui tahap Preprocessing, Total data yang digunakan berjumlah 600 komentar. Dari jumlah tersebut, sebanyak 538 komentar masuk dalam kategori sentimen positif, sementara 62 komentar lainnya tergolong ke dalam sentimen negatif.



Gambar 4. Hasil Pelabelan Bioskop Cinépolis Java Mall

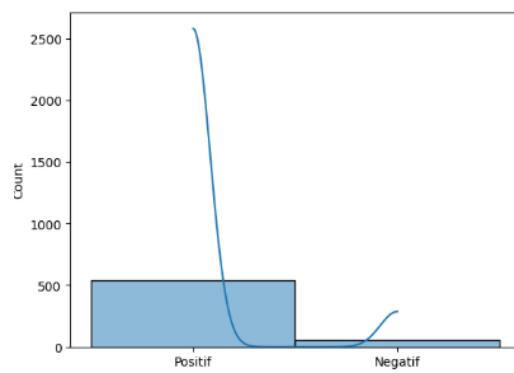
Gambar 4 memperlihatkan hasil pelabelan komentar dari Bioskop Cinépolis Java Mall setelah melalui tahap Preprocessing, Total data yang digunakan berjumlah 600 komentar. Dari jumlah tersebut, sebanyak 558 komentar

masuk dalam kategori sentimen positif, sementara 42 komentar lainnya tergolong ke dalam sentimen negatif.



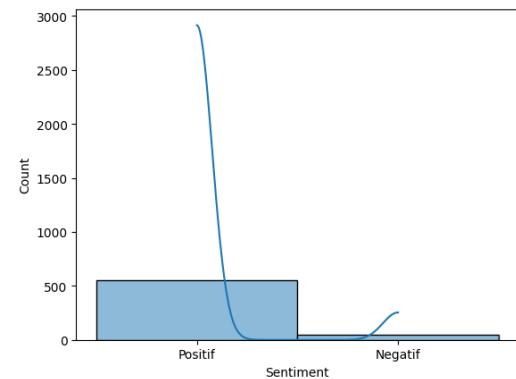
Gambar 5. Hasil Pelabelan Bioskop PARAGONXXI

Gambar 5 memperlihatkan hasil pelabelan komentar dari Bioskop PARAGONXXI setelah melalui tahap Preprocessing, Total data yang digunakan berjumlah 600 komentar. Dari jumlah tersebut, sebanyak 544 komentar masuk dalam kategori sentimen positif, sementara 56 komentar lainnya tergolong ke dalam sentimen negatif.



Gambar 6. Hasil Pelabelan Bioskop XXI Uptown Mall

Gambar 6 memperlihatkan hasil pelabelan komentar dari Bioskop XXI Uptown Mall setelah melalui tahap Preprocessing, Total data yang digunakan berjumlah 600 komentar. Dari jumlah tersebut, sebanyak 540 komentar masuk dalam kategori sentimen positif, sementara 60 komentar lainnya tergolong ke dalam sentimen negatif.



Gambar 7. Hasil pelabelan Bioskop XXI DP Mall

Gambar 7 memperlihatkan hasil pelabelan komentar dari Bioskop XXI DP Mall setelah melalui tahap Preprocessing, Total data yang digunakan berjumlah 600 komentar. Dari jumlah tersebut, sebanyak 552 komentar masuk dalam kategori sentimen positif, sementara 48 komentar lainnya tergolong ke dalam sentimen negatif.

D. Pengujian Metode Naïve Bayes Classifier

Pengujian dilakukan terhadap 120 data data test ulasan pelanggan yang berkaitan dengan pelayanan bioskop di semarang. Dalam proses ini, Algoritma Naive Bayes Classifier digunakan untuk mengklasifikasikan ulasan ke dalam dua kategori sentimen, yaitu Negatif dan Positif. Setelah klasifikasi dilakukan, hasil prediksi dievaluasi menggunakan pendekatan confusion matrix. Evaluasi ini mencakup pengukuran terhadap tiga metrik utama, yaitu precision, recall, dan accuracy. Untuk mengevaluasi seberapa baik model model dalam mengklasifikasikan sentimen dari tiap ulasan.

TABEL 8
HASIL KLASIFIKASI CITRAXXI

	Precision	Recall	F1-score	Support
Negative	0.00	0.00	0.00	12
Positive	0.90	1.00	0.95	108
Accuracy			0.90	120
Macro avg	0.45	0.50	0.47	120
Weighted avg	0.81	0.90	0.85	120

Hasil Klasifikasi dari bioskop CitraXXI Menunjukkan Untuk label positif memiliki nilai presisi 0.90, nilai recall 1.00, dan nilai F1-Score 0.95. Sedangkan label negatif memiliki nilai presisi 0.00, nilai recall 0.00, dan nilai F1-Score 0.00. Akurasi model ini secara keseluruhan adalah 0.90 dengan Rasio data training dan data testing adalah 80% untuk pelatihan dan 20% untuk pengujian.

TABEL 9
HASIL KLASIFIKASI BIOSKOP CINÉPOLIS JAVA MALL

	Precision	Recall	F1-score	Support
Negative	1.00	0.12	0.22	8
Positive	0.94	1.00	0.97	112
Accuracy			0.94	120
Macro avg	0.97	0.56	0.60	120
Weighted avg	0.95	0.94	0.92	120

Hasil Klasifikasi dari bioskop Cinépolis Java Mall Menunjukan Untuk label positif memiliki nilai presisi 0.94, nilai recall 1.00, dan nilai F1-Score 0.97. Sedangkan label negatif memiliki nilai presisi 1.00, nilai recall 0.12, dan nilai F1-Score 0.22. Akurasi model ini secara keseluruhan adalah 0.94 dengan Rasio data training dan data testing adalah 80% untuk pelatihan dan 20% untuk pengujian.

TABEL 10
HASIL KLASIFIKASI BIOSKOP PARAGONXXI

	Precision	Recall	F1-score	Support
Negative	0.00	0.00	0.00	11
Positive	0.91	1.00	0.95	109
Accuracy			0.91	120
Macro avg	0.45	0.50	0.48	120
Weighted avg	0.83	0.91	0.86	120

Hasil Klasifikasi dari bioskop ParagonXXI Menunjukan Untuk label positif memiliki nilai presisi 0.91, nilai recall 1.00, dan nilai F1-Score 0.95. Sedangkan label negatif memiliki nilai presisi 0.00, nilai recall 0.00, dan nilai F1-Score 0.00. Akurasi model ini secara keseluruhan adalah 0.91 dengan Rasio data training dan data testing adalah 80% untuk pelatihan dan 20% untuk pengujian.

TABEL 11
HASIL KLASIFIKASI BIOSKOP XXI UPTOWN MALL

	Precision	Recall	F1-score	Support
Negative	0.00	0.00	0.00	12
Positive	0.90	1.00	0.95	108
Accuracy			0.90	120
Macro avg	0.45	0.50	0.47	120
Weighted avg	0.81	0.90	0.85	120

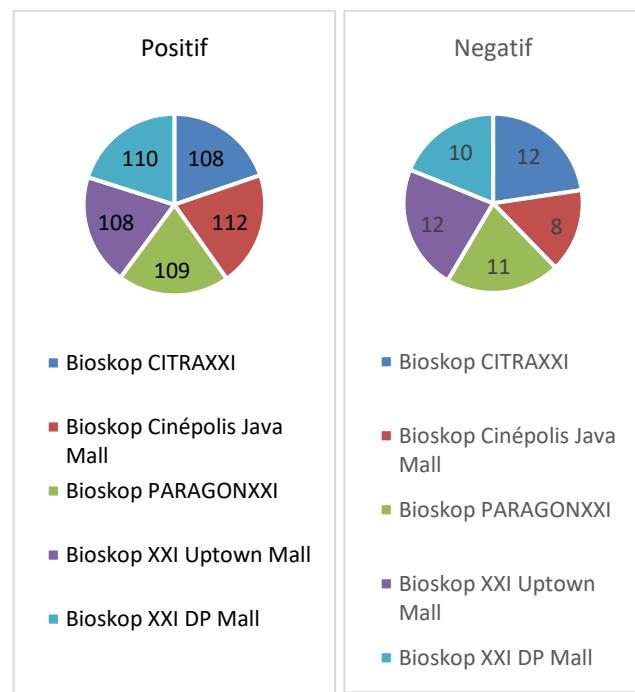
Hasil Klasifikasi dari bioskop XXI Uptown Mall Menunjukan Untuk label positif memiliki nilai presisi 0.90, nilai recall 1.00, dan nilai F1-Score 0.95. Sedangkan label negatif memiliki nilai presisi 0.00, nilai recall 0.00, dan nilai F1-Score 0.00. Akurasi model ini secara keseluruhan adalah 0.90 dengan Rasio data training dan data testing adalah 80% untuk pelatihan dan 20% untuk pengujian.

TABEL 12
HASIL KLASIFIKASI BIOSKOP XXI DP MALL

	Precision	Recall	F1-score	Support
Negative	0.50	0.10	0.17	10
Positive	0.92	0.99	0.96	110
Accuracy			0.92	120
Macro avg	0.71	0.55	0.56	120
Weighted avg	0.89	0.92	0.89	120

Hasil Klasifikasi dari bioskop XXI DP Mall Menunjukan Untuk label positif memiliki nilai presisi 0.92, nilai recall 0.99, dan nilai F1-Score 0.96. Sedangkan label negatif

memiliki nilai presisi 0.50, nilai recall 0.10, dan nilai F1-Score 0.17. Akurasi model ini secara keseluruhan adalah 0.92 dengan Rasio data training dan data testing adalah 80% untuk pelatihan dan 20% untuk pengujian.



TABEL 10
HASIL KLASIFIKASI BIOSKOP

Bioskop	Precision	Recall	F1-score	Accuracy
CitraXXI	0.90	1.00	0.95	0.90
Cinépolis Java Mall	0.94	1.00	0.97	0.94
ParagonXXI	0.91	1.00	0.95	0.91
XXI Uptown Mall	0.90	1.00	0.95	0.90
XXI DP Mall	0.92	0.99	0.96	0.92

IV. KESIMPULAN

Penelitian ini bertujuan untuk menganalisis sentimen pelanggan terhadap pelayanan lima bioskop di Kota Semarang dengan memanfaatkan metode Naïve Bayes Classifier. Data yang digunakan didapatkan dari 600 ulasan pelanggan di Google, yang kemudian Dikelompokkan dalam dua bagian: 480 training data untuk (80%) untuk membangun model dan Data test sebanyak 120 data (20%) untuk menguji akurasi model.

Model klasifikasi yang telah dibangun diuji pada data ulasan dari lima bioskop berbeda, masing-masing sebanyak 120 data. Hasil klasifikasi naïve bayes menunjukkan tingkat akurasi yang berbeda-beda pada tiap bioskop masing-masing bioskop CitraXXI memperoleh akurasi sebesar 0.90, Cinépolis Java Mall memperoleh akurasi tertinggi sebesar

0.94, Paragon XXI memperoleh akurasi sebesar 0.91, XXI Uptown Mall memperoleh akurasi sebesar 0.90 dan XXI DP Mall memperoleh akurasi sebesar 0.92

Selain menunjukkan perbandingan akurasi klasifikasi antar-bioskop, penelitian ini juga mengungkap pola sentimen pelanggan yang lebih mendalam. Secara umum, sentimen positif didominasi oleh apresiasi terhadap keramahan dan ketanggapan staf serta kenyamanan ruangan bioskop. Hal ini memperlihatkan bahwa kualitas layanan manusiawi dan pengalaman menonton yang nyaman menjadi faktor utama kepuasan pelanggan.

Di sisi lain, sentimen negatif terutama muncul dari keluhan terhadap fasilitas fisik, seperti pendingin ruangan (AC) yang tidak optimal, gangguan renovasi, hingga kondisi ruangan yang kurang terawat. Dengan demikian, meskipun secara keseluruhan layanan bioskop di Semarang dinilai baik, pengelola masih perlu memberikan perhatian lebih pada aspek pemeliharaan fasilitas agar kepuasan pelanggan dapat terjaga secara berkelanjutan. Hasil yang diperoleh menunjukkan bahwa Naïve Bayes dapat mencapai kinerja optimal dalam mengklasifikasikan sentimen ulasan pelanggan di lima bioskop dengan tingkat review tertinggi terdapat pada bioskop Cinépolis Java Mall.

DAFTAR PUSTAKA

- [1] Ernianti Hasibuan and Elmo Allistair Heriyanto, “Analisis Sentimen Pada Ulasan Aplikasi Amazon Shopping Di Google Play Store Menggunakan Naïve Bayes Classifier,” *J. Tek. dan Sci.*, vol. 1, no. 3, pp. 13–24, 2022, doi: 10.56127/jts.v1i3.434.
- [2] R. S. Malla and Y. Tanjung, “Pengaruh Harga , Lokasi dan Kualitas Pelayanan Terhadap Kepuasan Pelanggan pada Bioskop Cinepolis Plaza Medan Fair Salah satu Bioskop di Indonesia yaitu Cinepolis sebelumnya Bioskop Cinepolis beroperasi dengan nama Cinemaxx (dibawah nama PT Cinemaxx Glob,” *J. Manaj. Bisnis Kewirausahaan*, vol. 2, no. 3, 2023, [Online]. Available: <https://jurnal-stiepари.ac.id/index.php/jumbiwira/article/view/1094>
- [3] E. G. Maulana, L. P. Martha, and I. T. Kusumaningtias, “Analisis Strategi Komunikasi Pemasaran Terpadu Bioskop Platinum Cineplex Cibinong Square dalam Menarik Minat Menonton di Tengah Maraknya Video Streaming,” *J. CommLine*, vol. 07, no. 02, pp. 94–105, 2022.
- [4] O. B. Laksawicaka, B. Setioko, and E. Setyowati, “Gedung Bioskop Di Kota Semarang (Penekanan Desain Arsitektur Post Modern),” *Imaji*, vol. 3, no. 4, pp. 537–544, 2014, [Online]. Available: <https://ejournal3.undip.ac.id/index.php/imaji/article/download/695/2/6676>
- [5] Fransiska Vina Sari and Arief Wibowo, “Analisis Sentimen Pelanggan Toko Online Jd.Id Menggunakan Metode Naïve Bayes Classifier Berbasis Konversi Ikon Emosi,” *J. SIMETRIS*, vol. 10, no. 2, pp. 681–686, 2019.
- [6] A. Afif, “Analisis Metode Klasifikasi Data Naïve Bayes dan SVM Dalam Menentukan Keunikan Hotel,” *JUTEKINF (Jurnal Teknol. Komput. dan Informasi)*, vol. 11, no. 1, pp. 38–46, 2023, doi: 10.52072/jutekinf.v1i1.526.
- [7] M. I. Tuku, E. Umar, and A. Adis, “Analisis Sentimen Review Hotel Menggunakan Algoritma Naïve Bayes Pada Ella Hotel Tambolaka,” vol. 02, no. 03, pp. 346–356, 2024.
- [8] A. Cahya Kamilla, N. Priyani, R. Priskila, and V. Handrianus Pranatawijaya, “Analisis Sentimen Film Agak Laen Dengan Kecerdasan Buatan: Text Mining Metode Naïve Bayes Classifier,” *JATI (Jurnal Mhs. Tek. Infom.*, vol. 8, no. 3, pp. 2923–2928, 2024, doi: 10.36040/jati.v8i3.9587.
- [9] F. S. Jumeilah, “Klasifikasi Opini Masyarakat Terhadap Jasa Ekspedisi JNE dengan Naïve Bayes,” *J. Sist. Inf. Bisnis*, vol. 8, no. 1, p. 92, 2018, doi: 10.21456/vol8iss1pp92-98.
- [10] A. P. Wijaya and H. A. Santoso, “Naïve Bayes Classification pada Klasifikasi Dokumen Untuk Identifikasi Konten E-Government Naïve Bayes Classification on Document Classification to Identify E-Government Content,” *J. Appl. Intell. Syst.*, vol. 1, no. 1, pp. 48–55, 2016.
- [11] N. Nurdin, M. Suhendri, Y. Afrilia, and R. Rizal, “Klasifikasi Karya Ilmiah (Tugas Akhir) Mahasiswa Menggunakan Metode Naïve Bayes Classifier (NBC),” *Sistemasi*, vol. 10, no. 2, p. 268, 2021, doi: 10.32520/stmsi.v10i2.1193.
- [12] S. M. Salsabila, A. Alim Murtopo, and N. Fadhilah, “Analisis Sentimen Pelanggan Tokopedia Menggunakan Metode Naïve Bayes Classifier,” *J. Minfo Polgan*, vol. 11, no. 2, pp. 30–35, 2022, doi: 10.33395/jmp.v11i2.11640.
- [13] Yudi Widhiyasyana, Transmissia Semiawan, Ilham Gibran Achmad Mudzakir, and Muhammad Randi Noor, “Penerapan Convolutional Long Short-Term Memory untuk Klasifikasi Teks Berita Bahasa Indonesia,” *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 10, no. 4, pp. 354–361, 2021, doi: 10.22146/jnteti.v10i4.2438.
- [14] J. S. Informasi, F. Sains, D. Teknologi, U. Katolik, and M. Charitas, “Penggunaan Machine Learning (ML) dan Natural Language Processing (NLP) untuk mendeteksi Sentimen Ancaman Siber,” vol. 7, no. 1, pp. 1–6, 2025.
- [15] M. N. M. Al-Faruq, S. Nur’aini, and M. H. Aufan, “Perancangan Ui/Ux Semarang Virtual Tourism Dengan Figma,” *Walisongo J. Inf. Technol.*, vol. 4, no. 1, pp. 43–52, 2022, doi: 10.21580/wjit.2022.4.1.12079.
- [16] N. Rahayu, “Analisis Sentimen Pasar Otomotif Mobil : Tweet Twitter Menggunakan Naïve Bayes,” vol. 8, no. 1, pp. 113–120, 2017.
- [17] N. Nurwanda, N. Suarna, and W. Prihartono, “Penerapan Nlp (Natural Language Processing) Dalam Analisis Sentimen Pengguna Telegram Di Playstore,” *JATI (Jurnal Mhs. Tek. Infom.*, vol. 8, no. 2, pp. 1841–1846, 2024, doi: 10.36040/jati.v8i2.8469.
- [18] D. Mustikasari, I. Widaningrum, R. Arifin, and W. H. E. Putri, “Comparison of Effectiveness of Stemming Algorithms in Indonesian Documents,” *Proc. 2nd Borobudur Int. Symp. Sci. Technol. (BIS-STE 2020)*, vol. 203, pp. 154–158, 2021, doi: 10.2991/aer.k.210810.025.
- [19] K. Adib, M. R. Handayani, W. D. Yuniar, and K. Umam, “Opini Publik Pasca-Pemilihan Presiden: Eksplorasi Analisis Sentimen Media Sosial X Menggunakan SVM,” *SINTECH (Science Inf. Technol. J.*, vol. 7, no. 2, pp. 80–91, 2024, doi: 10.31598/sintechjournal.v7i2.1581.
- [20] C. A. Ul Hassan, M. S. Khan, and M. A. Shah, “Comparison of machine learning algorithms in data classification,” *ICAC 2018 - 2018 24th IEEE Int. Conf. Autom. Comput. Improv. Product. through Autom. Comput.*, no. September, pp. 1–6, 2018, doi: 10.23919/IConAC.2018.8748995.