ANALISIS SENTIMEN KOMENTAR FILM HOROR "WAKTU MAGHRIB" MENGGUNAKAN METODE *MULTINOMIAL NAÏVE BAYES*

TUGAS AKHIR

Diajukan oleh:

ROSSI WULANDARI
NIM : 220705055

Mahasiswa Fakultas Sains dan Teknologi
Program Studi Teknologi Informasi

FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI AR-RANIRY
BANDA ACEH
2026 M/1447 H

LEMBAR PERSETUJUAN

ANALISIS SENTIMEN KOMENTAR FILM HOROR "WAKTU MAGHRIB" MENGGUNAKAN METODE *MULTINOMIAL NAÏVE BAYES*

TUGAS AKHIR

Diajukan Kepada Fakultas Sains dan Teknologi
Universitas Islam Negeri (UIN) Ar-Raniry Banda Aceh
Sebagai Salah Satu Beban Studi Memperoleh Gelar Sarjana (S1) dalam Prodi
Teknologi Informasi

Oleh:
ROSSI WULANDARI
NIM : 220705055
Mahasiswa Fakultas Sains dan Teknologi
Program Studi Teknologi Informasi

Disetujui untuk Dimunaqasyahkan Oleh:

Pembimbing I,	Pembimbing II,
Mulkan Fadhli, S.T.,M.T.	**Fathiah, M.Eng.,**
NIP. 198811282020121006	**NIP. 198606152019032010**

Mengetahui,
Ketua Program Studi Teknologi Informasi

Malahayati, M.T
NIP. 198301272015032003

LEMBAR PENGESAHAN

ANALISIS SENTIMEN KOMENTAR FILM HOROR "WAKTU MAGHRIB" MENGGUNAKAN METODE MULTINOMIAL NAÏVE BAYES

TUGAS AKHIR

Telah Diuji Oleh Panitia Ujian Munaqasyah Tugas Akhir
Fakultas Sains dan Teknologi UIN Ar-Raniry Banda Aceh dan Dinyatakan Lulus
Serta Diterima Sebagai Salah Satu Beban Studi Program Sarjana (S1)
Dalam Program Studi Teknologi Informasi

Pada Hari/Tanggal: Rabu, 12 Agustus 2026 M
28 Shafar 1448 H
Di Darussalam, Banda Aceh

Panitia Ujian Munaqasyah Tugas Akhir

Ketua,	Sekretaris,
Mulkan Fadhli, M.T	**Fathiah, M.Eng**
NIP. 198811282020121006	**NIP. 198606152019032010**
Penguji I,	Penguji II,
Dr.Hendri Ahmadian, M.I.M	**Ridha Ilahi, M.T**
NIP. 198301042014031002	**NIP. 197905302014031001**

Mengetahui:
Dekan Fakultas Sains dan Teknologi
UIN Ar-Raniry Banda Aceh

Prof. Dr. Ir. Muhammad Dirhamsyah, M.T., IPU
NIP. 196210021988111001

LEMBAR PERNYATAAN KEASLIAN TUGAS AKHIR

Yang bertanda tangan di bawah ini:

Nama	: Rossi Wulandari
Nim	: 220705055
Prodi	: Teknologi Informasi
Fakultas	: Sains dan Teknologi
Judul	: Analisis Sentimen Komentar Film Horor "Waktu Maghrib" Menggunakan Metode Multinomial Naïve Bayes

Dengan ini menyatakan bahwa dalam penulisan tugas akhir ini, saya:

1. Tidak menggunakan ide orang lain tanpa mampu mengembangkan dan mempertanggungjawabkan;
2. Tidak melakukan plagiasi terhadap naskah karya orang lain;
3. Tidak menggunakan karya orang lain tanpa menyebutkan sumber asli atau tanpa izin pemilik karya;
4. Tidak memanipulasi dan memalsukan data;
5. Mengerjakan sendiri karya ini dan mampu bertanggungjawab atas karya ini.

Bila dikemudian hari ada tuntutan dari pihak lain atas karya saya, dan telah melalui pembuktian yang dapat dipertanggungjawabkan dan ternyata memang ditemukan bukti bahwa saya telah melanggar pernyataan ini, maka saya siap dikenai sanksi berdasarkan aturan yang berlaku di Fakultas Sains dan Teknologi UIN Ar-Raniry Banda Aceh.

Demikian pernyataan ini saya buat dengan sesungguhnya dan tanpa paksaan dari pihak manapun.

Banda Aceh, 12 Agustus 2026
Yang Menyatakan

(Rossi Wulandari)

ABSTRAK

Perkembangan platform digital memungkinkan masyarakat menyampaikan opini terhadap film melalui kolom komentar YouTube. Jumlah komentar yang besar menyebabkan analisis secara manual menjadi kurang efisien, sehingga diperlukan metode otomatis untuk mengidentifikasi kecenderungan sentimen penonton. Penelitian ini bertujuan mengimplementasikan metode *Multinomial Naïve Bayes* untuk mengklasifikasikan sentimen komentar penonton film horor *Waktu Maghrib* ke dalam kategori positif, netral, dan negatif serta mengevaluasi performa model. Data diperoleh menggunakan YouTube Data API sebanyak 2.000 komentar. Setelah melalui tahap *text preprocessing* yang meliputi *cleaning*, *filtering* bahasa, *case folding*, *tokenizing*, *stopword removal*, normalisasi kata, dan *stemming*, diperoleh 1.087 komentar berbahasa Indonesia. Sebanyak 500 komentar dipilih secara acak dan diberi label secara manual sebagai data penelitian. Dataset kemudian dibagi dengan rasio 80:20 menjadi 400 data pelatihan dan 100 data pengujian. Ekstraksi fitur dilakukan menggunakan TF-IDF dengan unigram dan bigram, kemudian klasifikasi dilakukan menggunakan Multinomial Naïve Bayes. Evaluasi model menggunakan *confusion matrix*, *accuracy*, *precision*, *recall*, dan *F1-score*. Hasil penelitian menunjukkan bahwa metode Multinomial Naïve Bayes mampu mengklasifikasikan sentimen komentar penonton dengan akurasi sebesar 62%, sehingga dapat digunakan untuk memberikan gambaran umum mengenai respons penonton terhadap film *Waktu Maghrib*.

Kata Kunci: Analisis sentimen, *Multinomial Naïve Bayes*, TF-IDF, Youtube *Waktu Maghrib*.

ABSTRACT

The growth of digital platforms enables the public to express their opinions on films via YouTube comment sections. The sheer volume of comments renders manual analysis inefficient, necessitating automated methods to identify audience sentiment trends. This study aims to implement the *Multinomial Naïve Bayes* method to classify audience sentiment regarding the horror film ”Waktu Maghrib” into positive, neutral, and negative categories, and to evaluate the model's performance. Data consisting of 2,000 comments was collected using the YouTube Data API. Following text preprocessing comprising *cleaning*, language *filtering*, *case folding*, *tokenization*, *stopword removal*, word normalization, and *stemming* 1,087 Indonesian-language comments were obtained. A subset of 500 comments was randomly selected and manually labeled for the study. The dataset was then split in an 80:20 ratio into 400 training samples and 100 testing samples. Feature extraction was performed using TF-IDF with unigrams and bigrams, followed by classification using *Multinomial Naïve Bayes*. Model evaluation utilized a *confusion matrix*, *accuracy*, *precision*, *recall*, and *F1-score*. The results indicate that the *Multinomial Naïve Bayes* method can classify audience sentiment with an accuracy of 62%, thereby providing a general overview of audience response to the film “Waktu Maghrib”.

Keyword: Sentiment analysis, *Multinomial Naïve Bayes*, TF-IDF, YouTube, *Waktu Maghrib*.

KATA PENGANTAR

Alhamdulillahirabbil'alamin, segala puji serta syukur penulis panjatkan ke hadirat Allah SWT atas segala rahmat, nikmat, dan kemudahan yang telah diberikan sehingga penulis dapat menyelesaikan skripsi yang berjudul **"Analisis Sentimen Komentar Film Horor "Waktu Maghrib" Menggunakan Metode Multinomial Naïve Bayes"**. Shalawat serta salam semoga senantiasa tercurah kepada Nabi Muhammad SAW sebagai teladan dalam menuntut ilmu dan menjalani kehidupan. Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana pada Program Studi Teknologi Informasi, Fakultas Sains dan Teknologi, Universitas Islam Negeri (UIN) Ar-Raniry Banda Aceh.

Penulis menyadari bahwa dalam penyusunan skripsi ini masih terdapat kekurangan, baik dari segi penyajian maupun analisis. Hal tersebut tidak terlepas dari keterbatasan pengetahuan, kemampuan, dan pengalaman penulis. Oleh karena itu, penulis sangat mengharapkan kritik dan saran yang membangun dari berbagai pihak demi penyempurnaan skripsi ini.

Penulis tuturkan serta persembahkan pada :

1. Ayah tercinta, Ismidin, dan Ibu tercinta, Ermiati, yang menjadi alasan terbesar penulis untuk terus bertahan dan menyelesaikan setiap proses perkuliahan hingga sampai pada tahap ini. Terima kasih atas setiap doa yang tidak pernah putus, kasih sayang yang tidak ternilai. Semoga Allah SWT senantiasa memberikan kesehatan, kebahagiaan, umur yang panjang, serta membalas segala kebaikan dan pengorbanan Ayah dan Ibu dengan keberkahan.

2. Terimakasih kepada saudara dan saudari saya terutama kepada abang Fadhil Rahmat, atas setiap bantuan, perhatian, nasihat, yang diberikan sejak awal perkuliahan hingga penulis dapat menyelesaikan skripsi ini. Dan kepada kakak saya Fitri Puspita Dewi, terima kasih atas

segala dukungan, perhatian, dan semangat yang selalu diberikan kepada penulis. Terakhir kepada kembaran saya Rossa Permata Sari, yang telah menemani setiap proses.

3. Bapak Mulkan Fadhli, M.T., selaku dosen pembimbing I yang telah meluangkan waktu memberikan arahan, masukan serta bimbingan.

4. Ibu Fathiah, M.Eng selaku dosen pembimbing II atas kesediaannya dalam memberikan bimbingan, masukan, serta dukungan kepada penulis selama proses penyusunan skripsi ini. Arahan yang diberikan sangat berarti dalam memperbaiki dan menyempurnakan penelitian ini.

5. Ibu Malahayati, M.T., selaku Ketua Program Studi Teknologi Informasi Fakultas Sains dan Teknologi Universitas Islam Negeri Ar-Raniry Banda Aceh.

6. Ibu Cut Ida Rahmadiana, S.Si, yang telah banyak membantu selama masa perkuliahan penulis.

7. Sahabat-sahabat terutama kepada dan seluruh angkatan 2022, terima kasih atas dukungan dan motivasi yang diberikan.

Banda Aceh, 12 Agustus 2026
Penulis,

Rossi Wulandari

DAFTAR ISI

DAFTAR GAMBAR

DAFTAR TABEL

DAFTAR LAMPIRAN

BAB I
PENDAHULUAN

1.1.Latar Belakang

Perkembangan teknologi informasi dan komunikasi yang semakin pesat telah mengubah cara masyarakat dalam mengakses dan menilai sebuah karya, termasuk film. Saat ini, penonton tidak hanya berperan sebagai konsumen pasif, tetapi juga aktif memberikan tanggapan, ulasan, dan opini melalui berbagai platform digital. Salah satu platform yang paling banyak digunakan adalah YouTube, di mana penonton dapat secara langsung memberikan komentar terhadap film yang mereka tonton, baik berupa trailer, cuplikan, maupun film secara keseluruhan. Komentar-komentar tersebut mencerminkan persepsi, emosi, dan penilaian penonton terhadap kualitas film.

Dalam penelitian ini, objek yang dikaji adalah film horor Indonesia berjudul "*Waktu Maghrib*". Film tersebut dipilih karena memperoleh perhatian yang cukup besar dari masyarakat. Hal ini terlihat dari pencapaian jumlah penonton yang telah menembus lebih dari 1 juta penonton pada Februari 2023 dan meningkat menjadi lebih dari 2 juta penonton pada Maret 2023. Pada akhir masa tayangnya, film "*Waktu Maghrib*" tercatat memperoleh sekitar 2,4 juta penonton. Capaian tersebut menunjukkan bahwa film ini memiliki jangkauan penonton yang cukup luas sehingga menghasilkan tanggapan dan opini dari masyarakat yang dapat dimanfaatkan sebagai sumber data penelitian.

Jumlah komentar yang terus meningkat membuat analisis secara manual menjadi tidak efektif akibat memakan waktu lama dan rentan terhadap subjektivitas. Oleh karena itu, dibutuhkan pendekatan otomatis yang efisien dalam mengolah data teks berskala besar. Analisis sentimen hadir sebagai solusi untuk mengklasifikasikan opini penonton ke dalam kategori positif, negatif, dan netral, sehingga mampu memberikan gambaran umum terkait kecenderungan respons publik terhadap film "*Waktu Maghrib*".

Platform YouTube dipilih sebagai sumber data dalam penelitian ini karena menyediakan data komentar melalui YouTube Data API yang dapat digunakan untuk memperoleh tanggapan pengguna terhadap film Waktu Maghrib. Dalam

penelitian ini, sebanyak 2.000 komentar diperoleh dari video terkait film Waktu Maghrib. Data hasil pengambilan kemudian melalui tahap pra-pemrosesan yang meliputi *cleaning*, *case folding*, *filtering* bahasa untuk mempertahankan komentar berbahasa Indonesia, *tokenizing*, *stopword removal*, normalisasi kata *slang*, dan *stemming*. Setelah proses *filtering* bahasa dan pembersihan data tersebut, diperoleh 1.087 komentar yang memenuhi kriteria penelitian. Selanjutnya, sebanyak 500 komentar dipilih secara acak untuk proses pelabelan manual ke dalam tiga kelas sentimen, yaitu positif, netral, dan negatif.

Pelabelan manual dipilih untuk meningkatkan validitas data karena mempertimbangkan konteks kalimat yang tidak selalu dapat dikenali secara tepat oleh metode pelabelan otomatis. Data yang telah diberi label kemudian dibagi menjadi data latih dan data uji. Selanjutnya, data diproses menggunakan pembobotan TF-IDF dan diklasifikasikan menggunakan metode *Multinomial Naïve Bayes*.

Penelitian mengenai analisis sentimen komentar film menggunakan berbagai algoritma klasifikasi telah banyak dilakukan. Namun, penelitian yang secara spesifik menganalisis sentimen komentar penonton film Waktu Maghrib menggunakan metode *Multinomial Naïve Bayes* masih terbilang terbatas. Dengan demikian, penelitian ini dilakukan untuk menganalisis sentimen komentar penonton terhadap film Waktu Maghrib menggunakan metode *Multinomial Naïve Bayes*.

Algoritma *Multinomial Naïve Bayes* diterapkan dalam klasifikasi sentimen karena sesuai untuk pengolahan data teks dengan representasi TF-IDF serta memiliki proses komputasi yang relatif efisien. Kinerja model dalam penelitian ini dievaluasi menggunakan *accuracy*, *precision*, *recall*, dan *F1-score*.

Berdasarkan uraian tersebut, penelitian ini dilakukan untuk menganalisis sentimen komentar penonton film horor "*Waktu Maghrib*" pada platform YouTube menggunakan metode *Multinomial Naïve Bayes*. Hasil penelitian diharapkan dapat memberikan gambaran mengenai kecenderungan sentimen penonton terhadap film "*Waktu Maghrib*".

1.2.Rumusan Masalah

1. Bagaimana penerapan metode *Multinomial Naïve Bayes* dalam mengklasifikasikan sentimen komentar penonton terhadap film horor *"Waktu Maghrib"* pada platform YouTube?
2. Bagaimana performa metode *Multinomial Naïve Bayes* dalam mengklasifikasikan sentimen komentar penonton terhadap film horor *"Waktu Maghrib"* berdasarkan nilai akurasi, presisi, *recall*, dan *F1-score*?

1.3.Tujuan Penelitian

Adapun tujuan penelitian ini adalah sebagai berikut:

1. Mengimplementasikan metode *Multinomial Naïve Bayes* untuk mengategorikan sentimen positif, negatif, serta netral dalam komentar penonton film horor *"Waktu Maghrib"* di platform YouTube.
2. Mengevaluasi performa metode *Multinomial Naïve Bayes* berdasarkan nilai akurasi, presisi, *recall*, dan *F1-score*.

1.4.Batasan Penelitian

Batasan penelitian ini mencakup beberapa hal berikut:

1. Data yang digunakan dalam penelitian ini berupa komentar penonton yang diperoleh dari platform YouTube pada video yang berkaitan dengan film horor *"Waktu Maghrib"*.
2. Data yang digunakan merupakan komentar berbahasa Indonesia yang diperoleh melalui YouTube Data API.
3. Penelitian ini menggunakan 500 komentar yang telah dipilih sebagai sampel dan diberi label sentimen secara manual ke dalam kelas positif, negatif, dan netral.
4. Metode klasifikasi yang digunakan dalam penelitian ini adalah *Multinomial Naïve Bayes* tanpa membandingkannya dengan algoritma klasifikasi lainnya.
5. Proses ekstraksi fitur dilakukan menggunakan metode TF-IDF (*Term Frequency–Inverse Document Frequency*)**.**

6. Hasil penelitian hanya menggambarkan analisis sentimen terhadap film horor *"Waktu Maghrib"*, sehingga tidak dapat digeneralisasikan untuk seluruh film horor di Indonesia.

1.5.Manfaat Penelitian

Adapun manfaat dalam penelitian ini adalah sebagai berikut:

1. Bagi Industri Perfilman / Pengelola Platform Digital:
    a. Membantu mengidentifikasi kecenderungan sentimen penonton terhadap film secara otomatis dan sistematis berdasarkan komentar pada platform YouTube.
    b. Menjadi bahan evaluasi bagi rumah produksi maupun pengelola platform digital dalam memahami respons penonton terhadap film yang dipublikasikan.
    c. Memberikan informasi yang dapat digunakan sebagai dasar dalam penyusunan strategi promosi, distribusi, dan pengembangan konten film sesuai dengan opini masyarakat.
2. Bagi Masyarakat (Penonton / Pengguna Platform Digital):
    a. Memberikan gambaran mengenai kecenderungan sentimen masyarakat terhadap film horor *"Waktu Maghrib"* berdasarkan hasil analisis komentar di YouTube
    b. Membantu masyarakat memperoleh informasi mengenai respons penonton terhadap film secara lebih objektif melalui hasil analisis data.
    c. Menjadi referensi bagi calon penonton dalam mengetahui persepsi publik terhadap film sebelum memutuskan untuk menontonnya.
3. Bagi Penulis
    a. Menambah wawasan dan pemahaman mengenai penerapan metode *Multinomial Naïve Bayes* dalam analisis sentimen berbasis *text mining*
    b. Meningkatkan keterampilan dalam proses pengumpulan

data, *preprocessing*, ekstraksi fitur menggunakan TF-IDF, penerapan *Random Oversampling*, serta evaluasi model klasifikasi.

c. Menjadi pengalaman dalam melakukan penelitian di bidang *machine learning*, *text mining*, dan *natural language processing*.

BAB II

TINJAUAN PUSTAKA

2.1 Penelitian Terdahulu

Dalam pengerjaan penelitian ini, dibutuhkan penelaahan terhadap sejumlah studi terdahulu yang telah dilaksanakan sebelumnya sebagai bentuk telaah ilmiah dan pembanding. Penelitian terdahulu menjadi dasar dalam memahami perkembangan metode, pendekatan, serta hasil yang telah dicapai dalam bidang analisis sentimen, khususnya yang memanfaatkan data media sosial dan teknik klasifikasi teks. Melalui peninjauan tersebut, peneliti dapat mengidentifikasi kesamaan, perbedaan, serta celah penelitian yang masih dapat dikembangkan. Oleh karena itu, beberapa penelitian yang relevan dengan topik yang diangkat dirangkum dan disajikan pada Tabel 2.1 untuk memperjelas posisi dan kontribusi penelitian ini.

Tabel 2. 1 Penelitian Terdahulu

Peneliti (Tahun)	Judul Penelitian	Hasil Penelitian	Persamaan dan Perbedaan
Nurpandi, (2024)	Analisis Sentimen Terhadap Kinerja Kepolisian Indonesia Menggunakan Metode *Multinomial Naive Bayes*, Long Short-Term Memory, dan	Penelitian membandingkan metode *Multinomial Naïve Bayes* dan LSTM pada data Twitter mengenai kinerja Kepolisian Indonesia. Model *Multinomial Naïve Bayes* memperoleh akurasi 78%, precision 78,5%, recall 77%, dan F1-score 77%,	Persamaan: Sama-sama menggunakan *Multinomial Naïve Bayes* untuk analisis sentimen pada data media sosial. Perbedaan: Penelitian terdahulu menggunakan data Twitter tentang kinerja Kepolisian Indonesia dan membandingkan beberapa metode, sedangkan penelitian ini menggunakan komentar YouTube film horor *Waktu*

	Lexicon-Based	sedangkan LSTM memperoleh akurasi 87%.	*Maghrib* dengan *Multinomial Naïve Bayes* sebagai metode utama.
(Firsttama et al., 2024)	Analisis Sentimen Komentar Youtube Konferensi Tingkat Tinggi G20 Menggunakan Metode *Naive Bayes*	Naïve Bayes mencapai akurasi 77% (precision 85%, recall 69%, F1-score 76%) dalam klasifikasi sentimen komentar YouTube terkait KTT G20 2022.	Persamaan: Menggunakan metode Naïve Bayes pada data komentar YouTube dengan tahapan *preprocessing*. Perbedaan: Objek data (KTT G20 vs film horor) dan jumlah kelas sentimen (2 vs 3).
Saragih & Kurniawan (2025)	Analisis Sentimen Menggunakan Metode *Naïve Bayes* Tentang Program Mudik Gratis Pemerintah Kota Medan 2024	Distribusi sentimen: 28% positif, 27% negatif, 45% netral; akurasi 51,67% dengan performa rendah pada kelas negatif (recall 12%).	Persamaan: Klasifikasi 3 kelas (positif/negatif/netral) pakai *Naive Bayes*. Perbedaan: Data komentar Instagram bukan khusus film.

(Iedwan et al., 2024)	Comparative Performance of SVM and Multinomial Naïve Bayes in Sentiment Analysis of the Film "Dirty Vote"	Akurasi yang diperoleh model SVM mencapai 97,27%, sementara *Multinomial Naïve Bayes* menghasilkan akurasi sebesar 96,70%.	Persamaan: Sama-sama menggunakan komentar YouTube, TF-IDF, dan Multinomial Naive Bayes. Perbedaan: Penelitian terdahulu membandingkan SVM dan Multinomial Naive Bayes pada film *Dirty Vote*, sedangkan penelitian ini hanya menggunakan Multinomial Naive Bayes pada film *Waktu Maghrib*.
(Rizki et al., 2024)	*Multinomial Naive Bayes* Algorithm for Indonesian Language Sentiment Classification Related to Jakarta International Stadium (JIS)	Model *Multinomial Naïve Bayes* berhasil mengklasifikasikan sentimen ulasan Google Maps tentang Jakarta International Stadium (JIS) dengan akurasi 83% menggunakan 2.971 data ulasan yang melalui tahapan *preprocessing* (case folding, tokenizing, stopword removal, dan stemming).	Persamaan: Menggunakan *Multinomial Naïve Bayes* untuk analisis sentimen dengan tahapan *preprocessing* dan pembobotan TF-IDF. Perbedaan : Penelitian ini menggunakan ulasan Google Maps Jakarta International Stadium (JIS), sedangkan penelitian ini menggunakan komentar YouTube film horor *Waktu Maghrib* dengan tiga kelas sentimen (positif, netral, dan negatif).

2.2 Landasan Teori

2.2.1 Analisis Sentimen

Analisis sentimen merupakan kajian berbasis komputer yang digunakan untuk mengidentifikasi opini, perasaan, dan emosi yang disampaikan dalam bentuk teks. Tujuan dari *opinion mining* adalah untuk menentukan apakah komentar dalam sekumpulan dokumen yang membahas suatu objek memiliki kecenderungan positif atau negatif (Salim & Syafrullah, 2023).

2.2.2. Text Mining

Text mining merupakan metode penemuan pengetahuan dalam dokumen teks yang mampu mengungkap wawasan berharga dari data tekstual yang tersedia. Proses ini menghadirkan tantangan tersendiri dalam mengeksplorasi informasi yang akurat demi membantu pengguna menemukan data yang mereka butuhkan. Agar lebih efektif, penerapan penemuan pengetahuan ini dapat diperbarui polanya secara berkala untuk kemudian diimplementasikan kembali ke dalam text mining. Secara umum, tahapan dalam *text mining* meliputi beberapa proses utama, yaitu *tokenizing*, *filtering*, *stemming*, tagging, dan analyzing (Y Ayu Putri Gabriella S, 2023).

Secara teknis, *Text mining* adalah proses menganalisis data teks menggunakan beragam teknik untuk mendeteksi pola maupun informasi penting. Tujuan proses ini yaitu mentransformasi data tidak terstruktur menjadi informasi yang lebih bermanfaat bagi kebutuhan tertentu. Di samping itu, *text mining* bermakna sebagai proses penarikan informasi dari berbagai dokumen teks. Fokus utamanya adalah mengidentifikasi kata-kata representatif dari isi dokumen agar hubungan antardokumen dapat dianalisis secara lebih mendalam. (Nitha Kumala Dewi, 2023).

2.2.3. Komentar Film

Komentar film merupakan bentuk opini atau tanggapan yang disampaikan oleh penonton terhadap suatu film setelah menonton, baik melalui platform online seperti IMDb, media sosial, maupun situs berbagi

video. Komentar tersebut berisi penilaian subjektif yang mencerminkan persepsi penonton terhadap kualitas film, seperti alur cerita, akting, dan aspek visual. Seiring meningkatnya penggunaan internet, jumlah komentar film juga semakin banyak dan beragam, sehingga sulit untuk dianalisis secara manual. Oleh sebab itu, penerapan analisis sentimen menjadi krusial untuk mengelompokkan tanggapan penonton ke dalam kelas sentimen positif dan negatif guna menghasilkan informasi yang terstruktur serta memberikan wawasan yang bermanfaat (Alwan & Ridla, 2024).

2.2.4. *Machine Learning*

Machine Learning (ML) merupakan cabang dari kecerdasan buatan (*artificial intelligence*) yang berfokus pada pengembangan sistem agar mampu belajar secara mandiri berdasarkan pemrosesan data. Komputer tidak perlu diprogram secara kaku untuk tugas tertentu, melainkan dilatih mengenali pola agar dapat mengambil keputusan berdasarkan hasil analisis data tersebut (Ratnasari, 2024).

Sebagai salah satu cabang utama Kecerdasan Buatan (AI), *machine learning* menjadi metode penting untuk mewujudkan mesin yang cerdas layaknya manusia. Melalui ML, komputer mampu mengenali pola data, memprediksi sesuatu, dan mengambil keputusan sendiri tanpa harus diprogram ulang untuk tiap tugas. Hal ini membuat AI dapat berkembang lebih cepat serta lebih fleksibel dalam menyelesaikan berbagai masalah (Ratnasari, 2024).

2.2.5. *Naïve Bayes*

Metode *Naïve Bayes* merupakan salah satu teknik pengklasifikasian yang mampu menentukan parameter yang dibutuhkan secara optimal meskipun hanya menggunakan data latih yang relatif terbatas. Metode ini memanfaatkan sejumlah atribut yang digunakan sebagai dasar dalam menentukan suatu keputusan, meskipun antar atribut tersebut dapat saling berkaitan (Pebdika et al., 2023).

Naïve Bayes merupakan teknik klasifikasi berbasis peluang sederhana yang

mengimplementasikan Teorema Bayes dengan asumsi bahwa setiap fitur bersifat independen. Kelebihan metode ini adalah efektivitasnya dalam menentukan parameter klasifikasi walau hanya didukung oleh jumlah sampel data latih yang terbatas (Pebdika et al., 2023). Persamaan dari teorema Bayes adalah :

$$P(H|X) = \frac{P(X|H).P(H)}{P(X)}$$

Di mana :

X :Data dengan class yang tidak dikenal.

H :Hipotesis data merupakan suatu class spesifik.

P(H|X) :Probabilitas hipotesis H berdasarkan kondisi X (posteriori probabilitas).

P(H) :Probabilitas hipotesis H (prior probabilitas).

P(X|H) :Probabilitas X berdasarkan kondisi pada hipotesis H

P(X) :Probabilitas X

Untuk mengimplementasikan metode *Naïve Bayes*, proses klasifikasi membutuhkan sejumlah atribut atau fitur sebagai petunjuk dalam menentukan kelas yang paling sesuai bagi sampel data yang dianalisis. Oleh karena itu, formulasi *Naïve Bayes* disesuaikan menjadi:

$$P(C|F1 \ldots Fn) = \frac{P(C)P(F1 \ldots Fn|C)}{P(F1 \ldots Fn)}$$

Dalam konteks penelitian ini, variabel C merepresentasikan kelas sentimen, yaitu positif, netral, dan negatif. Sedangkan F_1,F_2,…F_N merepresentasikan fitur berupa kata-kata yang diperoleh dari hasil pembobotan menggunakan metode *TF-IDF*. Setiap fitur mengindikasikan tingkat signifikansi suatu kata dalam sebuah dokumen, yang selanjutnya dimanfaatkan untuk mengestimasi probabilitas kepemilikan dokumen tersebut terhadap kategori sentimen tertentu menggunakan algoritma *Naïve Bayes*.

Di mana Variabel C merepresentasikan kelas, sementara variabel F1 ... Fn merepresentasikan set fitur yang digunakan sebagai kriteria klasifikasi. Dengan demikian, rumus tersebut menjelaskan bahwa probabilitas posterior yaitu

peluang sebuah sampel berfitur tertentu tergolong ke dalam kelas C diperoleh dari hasil perkalian antara probabilitas prior (peluang awal kelas C) dan likelihood (peluang kemunculan fitur pada kelas C), yang kemudian dibagi dengan *evidence* (probabilitas kemunculan fitur secara global) (Prasetiyowati & Sibaroni, 2024).

2.2.6. Multinomial Naïve Bayes

Multinomial Naïve Bayes merupakan algoritma probabilistik yang berpijak pada asumsi independensi antar-fitur. Kendati asumsi tersebut jarang terpenuhi dalam realitas, metode ini terbukti sangat efektif untuk kategorisasi teks. Penentuan kelas dokumen dilakukan dengan mengalkulasi probabilitas prior serta probabilitas bersyarat dari setiap token yang terkandung di dalam dokumen tersebut (Fitrianti & Yudhistira, 2025).

Multinomial Naïve Bayes dipilih karena kepopuleran metode ini didasari oleh kesederhanaannya dalam menangani data teks berorientasi frekuensi kemunculan kata. Pada dataset berukuran besar dengan variasi fitur yang kompleks seperti ulasan penggunan, *Multinomial Naïve Bayes* menawarkan efisiensi komputasi melalui perhitungan probabilitas bersyarat secara langsung. Lebih lanjut, berbagai literatur terdahulu mencatat bahwa penerapan metode ini pada analisis sentimen teks berbahasa Indonesia memberikan performa akurasi yang optimal (Suhendra dkk., 2024).

Rumus perhitungan *Multinomial Naïve Bayes* merupakan metode *supervised learning* sehingga setiap data perlu dilabeli sebelum dilakukan training, dapat digunakan persamaan 1 (Fitrianti & Yudhistira, 2025).

$$P(c|d) \propto P(c) \prod_{k=1}^{n} P\,(t_k|c)$$

keterangan:

P(c|d) : Probabilitas dokumen berada di kelas c

P(c) : Prior probabilitas suatu dokumen berada di kelas c

2.2.7. Rumus Slovin

Rumus Solvin adalah salah satu teori penarikan sampel yang paling populer untuk penelitian kuantitatif. Rumus Slovin biasa digunakan untuk pengambilan jumlah sampel yang harus representatif agar hasil penelitian dapat digeneralisasikan dan perhitungannya pun tidak memerlukan tabel jumlah sampel (Tunru et al., 2022).

$$\boldsymbol{n = \frac{N}{1 + N(e)^2}}$$

Keterangan:

n = Ukuran sampel/jumlah responden

N = Ukuran populasi

E = Persentase kelonggaran ketelitian

2.2.8. Text Preprocessing

Tahapan awal dalam proses *text mining* adalah *text preprocessing*. Tahap ini merupakan langkah penting yang perlu dilakukan sebelum data digunakan dalam pembangunan model. Proses *preprocessing* mencakup persiapan teks serta pembersihan data, yang menjadi bagian dari pengolahan dan analisis data. Tahapan ini melibatkan pemahaman terhadap karakteristik kumpulan data yang digunakan serta perencanaan teknik *preprocessing* yang sesuai (Hermawan & Jowensen, 2023). Secara umum, proses yang dilakukan dalam tahap preprocessing meliputi beberapa langkah sebagai berikut:

1. *Cleaning*

   Cleaning digunakan untuk menyeleksi dan menghapus elemen non-teks, berupa angka, *username* (@, RT, retweet, link, *hastag* (#), *HTML*, emoticon dan tanda baca lainnya seperti “,!$%&*”.

2. *Case Folding*

   Case folding berfungsi untuk menstandardisasi penulisan huruf di dalam data teks. Langkah ini penting untuk meningkatkan efisiensi analisis, mengingat teks sumber umumnya memiliki variasi penggunaan huruf

kapital yang tidak seragam. Dengan demikian, *case folding* mengubah seluruh teks menjadi huruf kecil (*lowercase*) agar memiliki format yang seragam.

3. *Tokenizing*

Tokenizing bertujuan untuk menguraikan teks berformat kalimat menjadi satuan kata tunggal (token) guna memfasilitasi pemrosesan pada tahapan selanjutnya (Astuti, 2024).

4. *Filtering* (*Stopword*)

Tahap filtrasi berfungsi untuk menyaring hasil tokenisasi agar hanya menyisakan kata-kata yang memiliki bobot informasi penting. Langkah ini diimplementasikan menggunakan teknik *stoplist* (penghapusan kata) atau *wordlist* (penyimpanan kata). *Stopword* yang dihapus mencakup kata fungsional atau kata sambung (seperti "yang", "dan", "di", "dari") karena tergolong non-deskriptif dan tidak esensial dalam analisis teks.

5. Normalisasi Kata

Normalisasi kata merupakan proses mengubah kata tidak baku, singkatan, kata slang, atau variasi penulisan menjadi bentuk kata yang lebih baku dan seragam agar dapat diproses secara lebih konsisten dalam analisis teks. Menurut Nugraha dan Rizqullah (2019), normalisasi diperlukan untuk menangani penggunaan bahasa informal dalam teks sehingga kata yang tidak terdapat dalam bentuk baku dapat disesuaikan ke bentuk yang sesuai. Dalam penelitian ini, normalisasi kata dilakukan terhadap komentar YouTube, misalnya kata "gk", "ga", dan "nggak" diubah menjadi "tidak", serta "bgt" menjadi "banget". Proses tersebut bertujuan mengurangi variasi kata yang memiliki makna sama sehingga representasi fitur pada tahap selanjutnya menjadi lebih konsisten.

6. *Stemming*

Stemming adalah proses mengubah suatu kata menjadi bentuk dasarnya. Dalam Bahasa Indonesia, proses ini digunakan untuk menghilangkan imbuhan yang melekat pada kata dasar. Proses *stemming* Bahasa Indonesia yang digunakan mengacu pada *library* Sastrawi yang

tersedia untuk Bahasa Pemrograman Python.

2.2.9. Ekstraksi Fitur

Ekstraksi fitur merupakan tahapan untuk mengidentifikasi dan mengkuantifikasi informasi kunci dari data mentah. Proses ini bertujuan mentransformasi representasi data berdimensi tinggi menjadi bentuk yang lebih sederhana dan terstruktur, tanpa mengeliminasi karakteristik esensial yang dibutuhkan dalam tahap klasifikasi maupun pengambilan keputusan (Wibisono et al., 2025).

Ekstraksi fitur merupakan tahap yang dilakukan setelah proses pra-pemrosesan (*preprocessing*). Tahap ini bertujuan untuk mengambil ciri atau fitur dari suatu objek citra yang dapat merepresentasikan karakteristik objek tersebut. Hasil dari ekstraksi fitur berupa nilai dalam bentuk matriks yang kemudian disimpan sebagai file dengan ekstensi .xml untuk dianalisis lebih lanjut pada tahap klasifikasi objek (Nur Cahyo et al., 2023).

Setelah itu, teks harus dikonversi menjadi representasi numerik yang dapat diproses algoritma pembelajaran mesin. Beberapa metode umum untuk mengekstraksi fitur dalam analisis sentimen adalah:

a. *Bag of Words*

Menghitung berapa kali kata muncul dalam setiap teks (Salman Al Markas et al., 2025).

b. *TF-IDF*

TF-IDF berfungsi untuk mengalokasikan nilai bobot numerik pada setiap kata di dalam dokumen teks. Melalui pendekatan ini, relevansi dan tingkat kepentingan suatu kata dapat diukur secara presisi berdasarkan kemunculannya dalam sebuah dokumen spesifik dibandingkan dengan kemunculannya pada kumpulan dokumen (dataset) (Wati et al., 2023). Secara matematis, rumus *TF-IDF* ialah sebagai berikut:

$$\text{tf-idf}(t,d) = \text{tf}(t,d) \times \text{idf}(t)$$

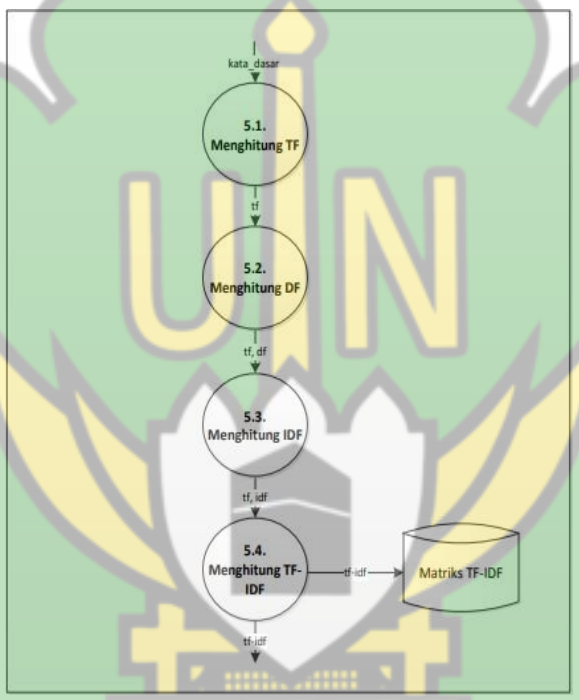
Dengan rumus tf(t,d) adalah sebagai berikut,

$$\mathrm{tf}(t,d) = \frac{f_{t,d}}{\sum_{t' \in d} f_{t',d}},$$

Selanjutnya, rumus idf(t) adalah sebagai berikut,

$$\mathrm{idf}(t) = \log \frac{1+n}{1+\mathrm{df}(t)} + 1,$$

Gambar arsitektur *TF-IDF* dapat dilihat pada Gambar 2.1 berikut:



Gambar 2. 1 Arsitektur TF-IDF

2.2.10. Evaluasi Model

a. Confusion Matrix

Pada tahap evaluasi, *Confusion Matrix* diterapkan untuk mengukur tingkat akurasi dan kinerja model klasifikasi. Matriks ini menyajikan matriks kontingensi yang membandingkan performa prediksi algoritma terhadap nilai kebenaran aktual (*ground truth*). Melalui tabel ini, distribusi

hasil klasifikasi dapat diidentifikasi berdasarkan kategori *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN). Struktur *Confusion Matrix* disajikan pada Tabel 2.2.

Tabel 2. 2 Confusion Matrix sumber (Palupi et al., 2023)

	Class. Actual Positive	**Class. True Negative**
Prediksi Class Positive	True Positive (TP)	False Negative (FN)
Prediksi Class Negative	False Positive (FP)	True Negative (TN)

b. Accuracy

Akurasi adalah tingkat rasio yang dapat dihitung dengan persamaan. Perhitungan nilai *accuracy* dapat dilakukan menggunakan persamaan berikut: (Palupi et al., 2023):

$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN}$$

c. Precision

Precision adalah mengukur tingkat ketepatan model dalam memprediksi suatu kelas sentimen tertentu. Sebagai contoh, dari seluruh ulasan yang diprediksi sebagai negatif, menunjukkan berapa banyak di antaranya yang benar-benar termasuk kategori negatif (Pambudi et al., 2024).

d. *Recall*

Recall yaitu mengukur kemampuan model dalam mengidentifikasi seluruh data yang benar-benar termasuk dalam suatu kelas tertentu. Sebagai contoh, menunjukkan seberapa banyak data aktual yang bernilai negatif dapat dikenali dengan benar oleh model (Pambudi et al., 2024).

e. F1-Score

F1-Score merupakan merupakan teknik untuk perhitungan yang menggabungkan *precision* dan *recall*. *F1-Score* dapat dihitung dengan persamaan (Palupi et al., 2023) :

2.2.11. Google Colab

Google Colab merupakan platform komputasi awan (*cloud-based platform*) yang menyediakan lingkungan eksekusi untuk menulis, menjalankan, serta membagikan kode program secara langsung terintegrasi dengan Google Drive. Platform ini mengusung antarmuka yang menyerupai Jupyter Notebook, tetapi menawarkan aksesibilitas yang lebih tinggi karena dapat dioperasikan melalui berbagai peramban (browser) tanpa memerlukan konfigurasi perangkat lunak lokal. Fleksibilitas ini memungkinkan proses pemrosesan data, pemodelan, hingga eksperimen komputasi dilakukan secara lebih efisien dan terstruktur (Muhammad Farid Dzakwan Wahyudi et al., 2024).

$$F1 - Score = 2\ x\ \frac{Presisi\ x\ Recall}{Presisi + Recall}$$

BAB III
METODOLOGI PENELITIAN

3.1. Jenis dan Pendekatan Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimen komputasi. Pendekatan kuantitatif digunakan karena data penelitian berupa komentar pengguna YouTube yang diolah dan direpresentasikan dalam bentuk numerik untuk selanjutnya diklasifikasikan berdasarkan kategori sentimen. Metode eksperimen komputasi digunakan karena penelitian melibatkan proses pengolahan data, pembentukan model klasifikasi, pengujian model, serta evaluasi performa model.

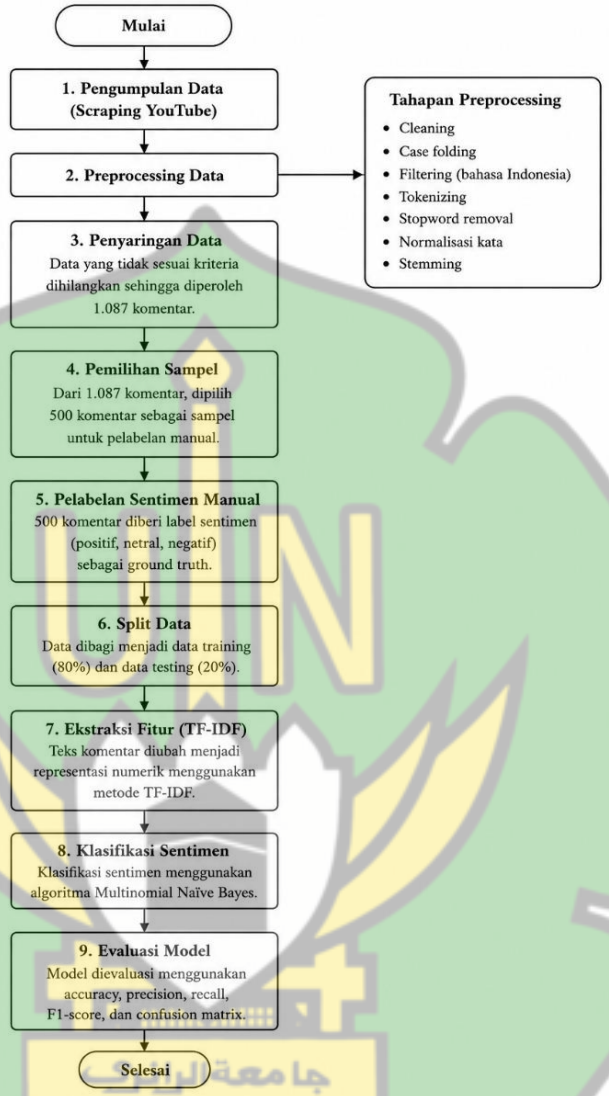
Penelitian ini menerapkan tahapan text mining yang meliputi pengumpulan data, *text preprocessing*, pelabelan sentimen, pembagian data menjadi data pelatihan (*training data*) dan data pengujian (*testing data*), ekstraksi fitur menggunakan *Term Frequency–Inverse Document Frequency* (TF-IDF), proses klasifikasi menggunakan algoritma *Multinomial Naïve Bayes*, serta evaluasi performa model.

Pendekatan eksperimen komputasi digunakan karena penelitian tidak hanya melakukan analisis terhadap data, tetapi juga membangun, melatih, dan menguji model klasifikasi menggunakan data pelatihan (*training data*) dan data pengujian (*testing data*). Hasil klasifikasi kemudian dievaluasi menggunakan *Confusion Matrix*, nilai *accuracy*, *precision*, *recall*, dan *F1-score* untuk mengetahui kemampuan model dalam mengklasifikasikan sentimen komentar penonton film *Waktu Maghrib* pada platform YouTube.

3.2. Tahapan Penelitian

Tahapan penelitian ini disusun secara sistematis berdasarkan alur kerja *machine learning* yang diimplementasikan menggunakan Google Colab dengan menerapkan algoritma *Multinomial Naïve Bayes*. Tahapan penelitian meliputi pengumpulan data, *text preprocessing*, pemilihan sampel data, pelabelan sentimen, pembagian data, ekstraksi fitur TF-IDF, klasifikasi menggunakan

Multinomial Naïve Bayes, dan evaluasi model. Alur metodologi secara menyeluruh disajikan pada Gambar 3.1.



Gambar 3. 1 Diagram alur penelitian

3.2.1. Pengumpulan Data

Tahap pertama penelitian adalah pengumpulan data komentar pengguna YouTube yang berkaitan dengan film horor "*Waktu Maghrib*". Pengambilan data dilakukan menggunakan YouTube Data API melalui program Python yang dijalankan pada lingkungan *Google Colab*.

Proses ini diawali dengan menentukan video ID sebagai parameter utama pengambilan komentar. Selanjutnya, sistem dihubungkan dengan API menggunakan *API Key* yang diperoleh dari Google Cloud Console. Komentar yang berhasil diambil kemudian disimpan dalam format CSV untuk memudahkan proses pengolahan data selanjutnya. Setelah data terkumpul, dilakukan pengecekan awal untuk menghapus komentar duplikat dan data kosong agar dataset lebih bersih dan siap diproses pada tahap berikutnya.

3.2.2. Text Preprocessing

Tahap kedua dalam penelitian ini adalah *text preprocessing*, yaitu proses pengolahan awal terhadap data komentar untuk mengurangi *noise* dan menghasilkan teks yang lebih terstruktur sebelum digunakan pada tahap pelabelan dan klasifikasi. Komentar YouTube memiliki bentuk penulisan yang beragam, seperti penggunaan huruf kapital, tanda baca, angka, URL, simbol, serta kata tidak baku atau kata *slang*. Oleh karena itu, tahapan *text preprocessing* dalam penelitian ini meliputi *cleaning*, *case folding*, *tokenizing*, *stopword removal*, normalisasi kata, dan *stemming*.

Langkah pertama adalah pembersihan (*cleaning*), yang mencakup penghapusan komponen-komponen yang tidak relevan dengan analisis sentimen, seperti angka, tanda baca, URL, dan simbol lainnya. Langkah berikutnya adalah *case folding*, yaitu mengubah seluruh karakter menjadi huruf kecil untuk menghindari perbedaan penafsiran terhadap istilah yang sama akibat variasi penggunaan huruf kapital.

Langkah selanjutnya adalah *filtering* (bahasa), yaitu proses penyeleksian komentar berdasarkan penggunaan bahasa. Pada tahap ini, komentar yang digunakan adalah komentar berbahasa Indonesia, sedangkan komentar yang menggunakan bahasa selain bahasa Indonesia tidak digunakan dalam proses analisis. Selanjutnya adalah *tokenisasi*, yaitu proses membagi kalimat atau teks menjadi unit-unit kata yang disebut token untuk mempermudah proses analisis dan pengolahan data.

Langkah selanjutnya adalah *stopword removal*, yaitu proses menghapus kata-kata umum yang sering muncul tetapi memiliki kontribusi yang relatif rendah terhadap analisis sentimen, seperti “dan”, “yang”, dan “di”. Selanjutnya dilakukan normalisasi kata, yaitu mengubah kata tidak baku atau kata slang ke dalam bentuk yang lebih sesuai dengan bahasa Indonesia agar makna kata tetap terjaga. Sebagai contoh, kata “serem” dinormalisasi menjadi “seram”, “bgt” menjadi “banget”, “yg” menjadi “yang”, dan “gk” menjadi “tidak”. Tahap normalisasi ini dilakukan sebelum *stemming* untuk mengurangi risiko perubahan bentuk kata slang yang dapat menghasilkan bentuk kata yang tidak sesuai dengan makna aslinya.

Tahap terakhir adalah *stemming*, yaitu proses mengubah kata menjadi bentuk dasarnya untuk menyeragamkan kata-kata yang memiliki akar kata yang sama. Pada penelitian ini, proses *stemming* bahasa Indonesia dilakukan menggunakan pustaka Sastrawi.

Hasil dari tahap ini teks telah dibersihkan dan siap untuk dikonversi menjadi bentuk numerik pada tahap ekstraksi fitur.

3.2.3 Pelabelan Sentimen

Setelah tahap text preprocessing, dilakukan pelabelan sentimen terhadap data komentar yang telah memenuhi kriteria penelitian. Dari 1.087 komentar hasil preprocessing, sebanyak 500 komentar dipilih sebagai sampel penelitian untuk dilakukan pelabelan secara manual. Pelabelan dilakukan berdasarkan makna dan kecenderungan opini yang terdapat dalam setiap komentar terhadap film horor “*Waktu Maghrib*”.

Proses pelabelan dilakukan secara manual (*ground truth*) dengan cara memberikan label sentimen secara langsung pada setiap komentar berdasarkan interpretasi peneliti terhadap makna yang terkandung dalam teks. Pelabelan ini mengacu pada tiga kategori sentimen, yaitu positif, negatif, dan netral.

Untuk menjaga konsistensi dalam proses pelabelan, setiap komentar dianalisis berdasarkan konteks dan kecenderungan opini yang disampaikan. Sentimen positif diberikan pada komentar yang menunjukkan apresiasi, kepuasan, atau penilaian positif terhadap film. Sentimen negatif diberikan pada komentar yang menunjukkan kritik, ketidakpuasan, atau penilaian negatif terhadap film. Sementara itu, sentimen netral diberikan pada komentar yang tidak menunjukkan kecenderungan positif maupun negatif secara jelas. Dengan demikian, setiap komentar dianalisis terlebih dahulu berdasarkan makna dan konteksnya sebelum diberikan label yang sesuai.

Hasil pelabelan sentimen selanjutnya digunakan sebagai variabel target (label) pada tahap pembagian data menjadi data training dan data testing. Label tersebut kemudian dimanfaatkan dalam proses pelatihan dan pengujian model klasifikasi menggunakan algoritma *Multinomial Naïve Bayes*.

3.2.4 Pembagian Data

Dataset yang telah melalui proses pelabelan sentimen kemudian dibagi menjadi data training dan data testing dengan perbandingan 80:20 menggunakan fungsi *train_test_split* dari pustaka *Scikit-learn*. Dari total 500 data berlabel, sebanyak 400 data (80%) digunakan sebagai data training dan 100 data (20%) digunakan sebagai data testing.

Pembagian data dilakukan dengan menggunakan parameter *stratify=y* untuk mempertahankan proporsi masing-masing kelas sentimen pada data training dan data testing. Dengan demikian, distribusi kelas positif, negatif, dan netral pada kedua subset data tetap mendekati proporsi pada dataset awal. Parameter *random_state=42* digunakan untuk memastikan proses pembagian data menghasilkan pembagian yang sama apabila penelitian dijalankan kembali dengan kondisi data yang sama.

Data *training* selanjutnya digunakan untuk proses pelatihan model

klasifikasi *Multinomial Naïve Bayes*, sedangkan data *testing* digunakan untuk menguji kemampuan model dalam mengklasifikasikan data yang belum digunakan selama proses pelatihan. Hasil klasifikasi pada data *testing* selanjutnya digunakan pada tahap evaluasi model.

3.2.5 Ekstraksi Fitur (TF-IDF)

Metode *Term Frequency–Inverse Document Frequency* (TF-IDF) digunakan untuk mengubah data teks menjadi representasi numerik yang dapat digunakan sebagai masukan dalam proses klasifikasi. Pada penelitian ini, ekstraksi fitur dilakukan setelah data dibagi menjadi data *training* dan data *testing* untuk mencegah terjadinya *data leakage*. Proses pembentukan *vocabulary* dan perhitungan bobot TF-IDF dilakukan hanya pada data training menggunakan fungsi *fit_transform()*. Selanjutnya, data testing diubah ke dalam representasi TF-IDF menggunakan fungsi *transform()* berdasarkan *vocabulary* yang telah dibentuk dari data training.

Ekstraksi fitur TF-IDF menggunakan parameter *ngram_range=(1,2)*, sehingga fitur yang dihasilkan terdiri atas *unigram* (satu kata) dan *bigram* (dua kata). Penggunaan *unigram* dan *bigram* bertujuan untuk mempertahankan informasi kata secara individual maupun kombinasi dua kata yang dapat memberikan konteks dalam analisis sentimen. Selain itu, parameter *min_df=2* digunakan untuk mengabaikan kata atau kombinasi kata yang hanya muncul pada satu dokumen sehingga fitur yang digunakan lebih relevan terhadap dataset.

Penerapan proses *fit_transform()* hanya pada data training dan *transform()* pada data testing dilakukan untuk mencegah *data leakage*, yaitu kondisi ketika informasi dari data testing ikut memengaruhi proses pembentukan fitur pada data training. Dengan demikian, data testing tidak digunakan dalam pembentukan *vocabulary* maupun perhitungan bobot TF-IDF selama proses pelatihan.

Hasil proses TF-IDF berupa matriks numerik yang merepresentasikan setiap komentar berdasarkan bobot kata yang terdapat di dalamnya. Matriks

TF-IDF dari data training digunakan sebagai masukan dalam proses pembentukan model *Multinomial Naïve Bayes*, sedangkan matriks TF-IDF dari data testing digunakan pada tahap pengujian model.

3.2.6 Proses Klasifikasi *Multinomial Naïve Bayes*

Pada tahap ini, metode *Multinomial Naïve Bayes* digunakan untuk melakukan klasifikasi sentimen terhadap komentar film horor *Waktu Maghrib*. Data *training* yang telah melalui proses pembagian data dan ekstraksi fitur TF-IDF digunakan untuk membangun model klasifikasi. Model mempelajari pola kemunculan kata atau fitur pada masing-masing dari tiga kategori sentimen, yaitu positif, netral, dan negatif.

Setelah proses pelatihan selesai, model digunakan untuk memprediksi kelas sentimen pada data *testing* yang tidak digunakan selama proses pelatihan. Hasil prediksi kemudian dibandingkan dengan label sebenarnya pada data *testing* untuk mengetahui kemampuan model dalam melakukan klasifikasi sentimen.

3.2.7 Evaluasi Model

Tahap evaluasi dilakukan untuk mengetahui kinerja model *Multinomial Naïve Bayes* dalam mengklasifikasikan komentar ke dalam tiga kategori sentimen, yaitu positif, netral, dan negatif. Evaluasi dilakukan berdasarkan hasil prediksi model terhadap data testing yang telah memiliki label sebenarnya (*ground truth*).

Performa model dievaluasi menggunakan *confusion matrix*, *accuracy*, *precision*, *recall*, dan *F1-score* dengan bantuan pustaka *Scikit-learn*. *Confusion matrix* digunakan untuk melihat jumlah prediksi yang benar dan salah pada masing-masing kelas sentimen, sedangkan *accuracy*, *precision*, *recall*, dan *F1-score* digunakan untuk mengukur kinerja model secara kuantitatif.

Evaluasi dilakukan untuk mengetahui kinerja model *Multinomial Naïve Bayes* berdasarkan hasil klasifikasi pada data *testing*. Hasil evaluasi

dianalisis berdasarkan nilai *accuracy*, *precision*, *recall*, dan *F1-score*, serta *confusion matrix* untuk mengetahui kemampuan model dalam mengklasifikasikan sentimen positif, netral, dan negatif.

3.3 Waktu dan Tempat Penelitian

Penelitian ini berlangsung pada rentang waktu Desember 2025 sampai Juli 2026. Dalam periode tersebut, dilaksanakan seluruh tahapan penelitian secara sistematis, mencakup pengumpulan data, pra pemrosesan teks (*text preprocessing*), pelabelan sentimen, pembagian data, ekstraksi fitur TF-IDF, klasifikasi menggunakan algoritma *Multinomial Naïve Bayes*, hingga evaluasi kinerja model.

Penelitian dilakukan secara daring dengan menggunakan Google Colab sebagai lingkungan komputasi berbasis *cloud*. Google Colab digunakan untuk menjalankan program Python dalam proses pengumpulan dan pengolahan data, *text preprocessing*, ekstraksi fitur TF-IDF, pembentukan dan pengujian model *Multinomial Naïve Bayes*, serta evaluasi hasil klasifikasi.

3.4 Sumber dan Pengumpulan Data

Data yang digunakan dalam penelitian ini bersumber dari platform YouTube, yaitu komentar pengguna pada video film horor "*Waktu Maghrib"* yang dijadikan sebagai objek penelitian. Kode Python yang dijalankan di *Google Colab* menggunakan YouTube Data API untuk mengambil data secara otomatis. Dengan menggunakan kunci API dari *Google Cloud Console*, sistem terhubung ke YouTube Data API setelah ID video ditentukan sebagai parameter utama untuk ekstraksi komentar. Melalui proses tersebut berhasil dikumpulkan sebanyak 2.000 komentar yang selanjutnya disimpan dalam format CSV sebagai dataset awal.

Dataset hasil *scraping* kemudian melalui tahap *text preprocessing* yang meliputi *cleaning*, *filtering* bahasa, *case folding*, *tokenizing*, *stopword removal*, normalisasi kata, dan *stemming*. Tahapan tersebut dilakukan untuk membersihkan dan menyeragamkan teks serta menyaring komentar yang tidak menggunakan bahasa Indonesia sehingga diperoleh data yang sesuai dengan kriteria penelitian. Setelah seluruh tahapan *text preprocessing* dan *filtering* bahasa dilakukan, diperoleh 1.087 komentar yang selanjutnya menjadi data hasil preprocessing.

Selanjutnya, 500 komentar dipilih secara acak dari total 1.087 komentar menggunakan fungsi *sample()* dengan parameter *random_state=42*. Sampel-sampel tersebut kemudian diberi label sentimen secara manual (*ground truth*) ke dalam tiga kategori: positif, netral, dan negatif. Dataset yang telah diberi label tersebut selanjutnya digunakan untuk pembagian data (*train-test split*), ekstraksi fitur menggunakan TF-IDF, klasifikasi menggunakan *Multinomial Naïve Bayes*, penerapan *Random Oversampling*, serta evaluasi model.

Untuk memberikan gambaran mengenai data yang digunakan, struktur dataset hasil pengambilan komentar disajikan pada Tabel 3.1

Tabel 3. 1 Struktur Dataset Awal

No	Nama Kolom	Tipe Data	Keterangan
1.	Komentar	Text	Komentar asli pengguna hasil *scraping* dari YouTube

Setelah proses pengumpulan data selesai, diperoleh sebanyak 2.000 komentar yang kemudian digunakan sebagai data awal penelitian. Komentar tersebut selanjutnya melalui tahapan preprocessing dan filtering sebelum dipilih sebagai sampel penelitian. Sebagai bukti proses pengambilan data secara otomatis, ditampilkan tangkapan layar (*screenshot*) hasil eksekusi program Python di Google Colab yang menunjukkan proses pengambilan komentar dari YouTube menggunakan YouTube Data API.

	Komentar
0	Rruttu
1	tadi saya tengok pada 19/4/2026
2	Jangan panggil aku namanya moy aku namanya moy...
3	Serem banget tinggal tulang doang hantu
4	<a href="https://www.youtube.com/watch?v=uFGE0...
5	keren
6	Serem banget aku Ampe tutup mata sama telinga ...
7	Serem banget ya Allah SWT
8	Si main kun anta 2 hipnotis 1.<a href="https:/...
9	jantung aku udah mau copoy

Gambar 3. 2 Hasil Pengambilan Data Komentar dari YouTube

Perangkat lunak Python yang digunakan untuk mengambil komentar dari YouTube menggunakan YouTube Data API ditampilkan pada Gambar 3.2. Sebelum tahap pra-pemrosesan teks dan pemrosesan tambahan, sebanyak 2.000 komentar berhasil dikumpulkan dan digunakan sebagai dataset awal.

Rincian jumlah data sebelum dan sesudah tahap *preprocessing* disajikan pada Tabel 3.2 berikut.

Tabel 3. 2 Jumlah Data Penelitian

Keterangan	Jumlah
Total komentar yang diperoleh	2000
Data setelah *preprocessing* dan *filtering*	1087
Data sampel untuk pelabelan manual	500
Data Training (80%)	400
Data Testing (20%)	100

Berdasarkan Tabel 3.2, proses pengambilan data menggunakan YouTube Data API menghasilkan sebanyak 2.000 komentar dari video film *Waktu Maghrib*. Selanjutnya, data melalui tahap *text preprocessing* dan *filtering* bahasa sehingga diperoleh sebanyak 1.087 komentar yang memenuhi kriteria penelitian. Sebanyak 500 komentar dipilih secara acak dari total tersebut untuk dijadikan sampel

penelitian. Komentar-komentar ini kemudian diklasifikasikan secara manual (*ground truth*) ke dalam tiga kategori: positif, netral, dan negatif. Setelah data berlabel dibagi menjadi set pelatihan dan pengujian dengan rasio 80:20, diperoleh 400 sampel pelatihan dan 100 sampel pengujian. Data *training* selanjutnya digunakan untuk proses pembentukan model *Multinomial Naïve Bayes*, sedangkan data *testing* digunakan untuk menguji kinerja model.

3.5 Populasi dan Sampel

Populasi dalam penelitian ini didasarkan pada jumlah penonton film *Waktu Maghrib* pada platform YouTube, yaitu sekitar 2,4 juta penonton. Untuk menentukan jumlah sampel yang digunakan dalam penelitian, dilakukan perhitungan menggunakan *Rumus Slovin* dengan tingkat kesalahan sebesar 5%. Rumus Slovin digunakan untuk menentukan jumlah minimum sampel yang dapat mewakili populasi penelitian. Adapun rumus yang digunakan adalah sebagai berikut:

$$\boldsymbol{n = \frac{N}{1 + N(e)^2}}$$

Keterangan:

- n = jumlah sampel
- N = jumlah populasi
- e = tingkat kesalahan (*margin of error*)

$$n = \frac{2.400.000}{1 + 2.400.000\ (0{,}05)^2}$$

$$n = \frac{2.400.000}{1 + 2.400.000\ (0{,}0025)^2}$$

$$n = \frac{2.400.000}{6.001}$$

$$n = 399{,}93$$

Berdasarkan hasil perhitungan tersebut, jumlah sampel minimum yang diperlukan adalah 399,93 atau dibulatkan menjadi 400 komentar. Jumlah tersebut tidak melebihi 500 komentar yang tersedia setelah proses penyaringan data. Oleh karena

itu, penelitian ini menggunakan 500 komentar sebagai sampel penelitian yang kemudian diberi label sentimen secara manual sebagai *ground truth*. Penggunaan 500 komentar tersebut telah memenuhi jumlah sampel minimum yang diperoleh berdasarkan *Rumus Slovin*.

3.6 Metode Multinomial Naïve Bayes

Salah satu pendekatan klasifikasi yang populer dalam analisis data teks adalah metode *Multinomial Naïve Bayes*, yang didasarkan pada *Teorema Bayes*. Pendekatan ini menggunakan frekuensi kemunculan kata dalam setiap kelas untuk menentukan kemungkinan suatu dokumen termasuk dalam kelas tertentu. Dalam penelitian ini, komentar YouTube diklasifikasikan ke dalam tiga kategori sentimen positif, netral, dan negatif menggunakan metode *Multinomial Naïve Bayes*.

Prinsip dasar algoritma ini adalah bahwa setiap kata atau fitur bersifat independen satu sama lain. Meskipun asumsi ini sederhana, *Multinomial Naïve Bayes* menunjukkan kinerja yang baik dalam mengategorikan data teks yang direpresentasikan menggunakan TF-IDF, serta menawarkan keunggulan berupa perhitungan yang cepat dan efisien.

Secara matematis, *Teorema Bayes* dinyatakan sebagai berikut

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

Keterangan:

a. P(C|X) = adalah probabilitas posterior suatu kelas setelah diketahui fitur.
b. P(X|C) = adalah probabilitas kemunculan fitur pada kelas tertentu (likelihood).
c. P(C) = adalah probabilitas awal (prior) suatu kelas.
d. P(X) = adalah probabilitas kemunculan fitur.

Data dalam penelitian ini diubah menjadi representasi numerik menggunakan metode TF-IDF setelah melalui tahap *text preprocessing*, pelabelan sentimen, dan pembagian data. Untuk mencegah *data leakage*, proses ekstraksi fitur dilakukan dengan menjalankan operasi *fit_transform()* pada data *training* dan operasi

transform() pada data *testing*. Ekstraksi fitur TF-IDF menggunakan *ngram_range=(1,2)* untuk menghasilkan fitur berupa unigram dan bigram, serta *min_df=2* untuk mengabaikan fitur yang muncul pada kurang dari dua dokumen.

Pemilihan *Multinomial Naïve Bayes* didasarkan pada kemampuannya dalam mengolah data teks, proses komputasi yang efisien, serta kesesuaiannya untuk melakukan klasifikasi dokumen berdasarkan frekuensi kemunculan fitur kata. Pada penelitian ini, *Multinomial Naïve Bayes* digunakan untuk mengklasifikasikan komentar ke dalam tiga kategori sentimen, yaitu positif, netral, dan negatif. Model dilatih menggunakan data *training* yang telah direpresentasikan dalam bentuk fitur TF-IDF, kemudian digunakan untuk memprediksi kelas sentimen pada data *testing*. Hasil klasifikasi selanjutnya dievaluasi menggunakan *confusion matrix*, *accuracy*, *precision*, *recall*, dan *F1-score* untuk mengetahui kinerja model.

3.7. Alat dan Bahan Penelitian

Dalam penelitian ini, alat dan bahan yang digunakan meliputi perangkat keras (*hardware*), perangkat lunak (*software*), dan data penelitian. Ketiga komponen tersebut berperan dalam mendukung proses pengumpulan data, pengolahan data, implementasi algoritma *Multinomial Naïve Bayes*, serta evaluasi hasil klasifikasi sentimen.

3.7.1. Perangkat Keras (*Hardware*)

Perangkat keras yang digunakan dalam penelitian ini berupa laptop atau komputer pribadi yang terhubung dengan jaringan internet. Perangkat tersebut digunakan untuk menjalankan program, mengakses *Google Colab*, serta memantau seluruh proses pengolahan data dan implementasi metode klasifikasi. Meskipun proses komputasi dilakukan secara daring (*cloud computing*), perangkat keras tetap berperan sebagai sarana utama dalam mendukung pelaksanaan penelitian. Adapun rincian perangkat keras yang digunakan disajikan pada Tabel 3.3.

Tabel 3. 3 Perangkat Keras (Hardware)

No	Perangkat	Spesifikasi/Keterangan
1	Laptop/PC	Acer

2	RAM	4 GB DDR4
3	Penyimpanan	1000 GB HDD
4	Jaringan internet	Koneksi internet untuk mengakses Google Colab, YouTube Data API, dan proses pengolahan data secara daring.

3.7.2 Perangkat Lunak (*Software*)

Seluruh tahapan penelitian didukung oleh perangkat lunak, termasuk pengumpulan data, pra-pemrosesan, ekstraksi fitur berbobot TF-IDF, klasifikasi menggunakan algoritma *Multinomial Naïve Bayes*, serta penilaian kinerja model. Tabel 3.4 menyajikan informasi mengenai perangkat lunak yang digunakan dalam penelitian ini.

Tabel 3. 4 Perangkat Lunak (Software)

No	Perangkat Lunak	Fungsi
1	*Google Colab*	Lingkungan eksekusi program berbasis *cloud*
2	Python	Bahasa pemrograman utama penelitian
3	Pandas	Pengolahan dan manipulasi dataset
4	NumPy	Operasi numerik dan pengolahan Array
5	Google API Client Library for Python	Mengambil komentar dari YouTube melalui API
6	*langdetect*	Deteksi bahasa pada komentar
7	NLTK	Melakukan proses *tokenizing*
8.	Sastrawi	Melakukan *stopword removal* dan *stemming* bahasa Indonesia.

9.	*Scikit-learn*	*TF-IDF*, *Naïve Bayes*, split data, dan evaluasi model
10	Matplotlib	Visualisasi data

3.7.3 Bahan Penelitian

Objek dasar dari metode analisis sentimen adalah data, yang menjadi materi penelitian. Sebuah dataset berisi komentar pengguna YouTube, yang dikumpulkan melalui teknik *data scraping* menggunakan YouTube Data API, menjadi sumber data penelitian ini. Data yang diperoleh kemudian melalui tahap *preprocessing* dan *filtering* sehingga menghasilkan data yang sesuai dengan kriteria penelitian. Dari data tersebut, dipilih sebanyak 500 komentar sebagai sampel penelitian yang selanjutnya dilakukan pelabelan sentimen secara manual ke dalam kategori positif, netral, dan negatif. Adapun rincian bahan penelitian yang digunakan disajikan pada Tabel 3.5 berikut.

Tabel 3. 5 Bahan penelitian

Bahan	**Sumber**	**Keterangan**
Dataset komentar video film *Waktu Maghrib*	YouTube	Data mentah hasil scraping sebanyak 2000 komentar menggunakan YouTube Data API
Dataset hasil *preprocessing* dan *filtering*	Hasil Pengolahan	Data komentar yang telah melalui preprocessing dan filtering sebanyak 1087 komentar
File dataset penelitian (data_500_label_manual.xlsx)	Hasil pengolahan	Dataset berisi 500 komentar yang dipilih secara acak dan digunakan sebagai sampel penelitian.
Label sentimen	Pelabelan manual	Label positif, netral, dan negatif (*ground truth*) yang diberikan secara manual.

BAB IV
HASIL DAN PEMBAHASAN

4.1. Hasil Penelitian

4.1.1. Hasil Pengumpulan Data

Pada tahap pertama penelitian, komentar pengguna mengenai film "*Waktu Maghrib*" dikumpulkan melalui platform YouTube. YouTube Data API digunakan untuk mengambil data dengan menggunakan bahasa pemrograman Python di lingkungan Google Colab.

Berdasarkan hasil proses *scraping*, diperoleh sebanyak 2.000 komentar. Tahapan penyaringan dan pra-pemrosesan kemudian digunakan untuk menyingkirkan komentar yang tidak relevan dengan penelitian, seperti duplikat, komentar kosong, komentar yang tidak menggunakan bahasa Indonesia, serta komentar yang tidak layak untuk diproses lebih lanjut.

Setelah proses penyaringan selesai dilakukan, jumlah data yang memenuhi kriteria penelitian menjadi 1.087 komentar. Berdasarkan jumlah populasi tersebut, perhitungan menggunakan rumus Slovin dengan tingkat kesalahan 5% menghasilkan jumlah sampel minimum sebanyak 293 komentar. Dalam penelitian ini, sebanyak 500 komentar dipilih sebagai sampel penelitian, yang jumlahnya telah melebihi batas minimum sampel berdasarkan perhitungan Slovin. Selanjutnya, 500 komentar tersebut dilakukan pelabelan sentimen secara manual sebagai *ground truth*. Data yang telah diberi label kemudian digunakan untuk proses pembagian data menjadi data *training* dan data *testing*, ekstraksi fitur menggunakan metode TF-IDF, serta klasifikasi sentimen menggunakan algoritma *Multinomial Naïve Bayes*.

Tingkat kesalahan sebesar 5% digunakan karena merupakan tingkat kesalahan yang umum digunakan dalam penelitian kuantitatif untuk menentukan batas toleransi kesalahan dalam pengambilan sampel. Penggunaan tingkat kesalahan tersebut memberikan tingkat kepercayaan sebesar 95%, sehingga sampel yang diperoleh diharapkan dapat memberikan representasi yang cukup terhadap populasi penelitian.

Keseluruhan data yang berhasil dikumpulkan disimpan ke dalam format *.csv* guna mengoptimalkan proses pemrosesan dan analisis pada tahap selanjutnya. Rincian mengenai kuantitas data penelitian ini dipaparkan pada Tabel 4.1.

Tabel 4. 1 Jumlah Data Penelitian

No	Keterangan	Jumlah
1	Total komentar hasil scraping YouTube	2000
2	Data setelah *preprocessing* dan *filtering*	1087
3	Data sampel untuk pelabelan manual (*ground truth*)	500

Setelah itu, data yang telah dikumpulkan ditampilkan sebagai *dataframe* untuk mempermudah tahapan pemrosesan selanjutnya. Gambar 4.1 menampilkan hasil akhir dari pengumpulan data tersebut.

	Komentar
0	Rruttu
1	tadi saya tengok pada 19/4/2026
2	Jangan panggil aku namanya moy aku namanya moy...
3	Serem banget tinggal tulang doang hantu 🤣
4	<a href="https://www.youtube.com/watch?v=uFGE0...
5	keren
6	Serem banget aku Ampe tutup mata sama telinga ...
7	Serem banget ya Allah SWT 😢😢😢
8	Si main kun anta 2 hipnotis 1.<a href="https:/...
9	jantung aku udah mau copoy
10	Serem bangettttttttttttttttttttttttttttt😲❤
11	c0mbi

Gambar 4. 1 Hasil Scraping Komentar YouTube

4.1.2. Hasil Text Preprocessing

Tahap pra pemrosesan teks (*text preprocessing*) diterapkan guna memurnikan serta menyiapkan data komentar sebelum ditransformasikan ke dalam proses klasifikasi sentimen. Pembersihan *noise* pada data bertujuan untuk meningkatkan relevansi informasi sehingga mempermudah pemodelan *machine learning*. Adapun eksekusi *text preprocessing* pada penelitian ini melibatkan enam langkah utama: *cleaning*, *case folding*, *filtering* bahasa, *tokenizing*,*stopword removal*, normalisasi kata dan *stemming*.

Pada tahap *cleaning*, berfokus pada pembersihan elemen yang kurang relevan, meliputi simbol, angka, tanda baca, emoji, tautan, dan karakter khusus lainnya yang tidak memiliki kontribusi terhadap analisis sentimen. Tahap selanjutnya berupa *case folding*, yaitu penyeragaman seluruh karakter menjadi huruf kecil guna mencegah timbulnya redundansi representasi kata akibat penggunaan huruf kapital.

Tahap berikutnya adalah *filtering* bahasa menggunakan deteksi bahasa untuk memastikan hanya komentar berbahasa Indonesia yang digunakan dalam penelitian. Tahap ini dilakukan untuk menghilangkan komentar berbahasa asing maupun komentar yang tidak dapat dikenali secara jelas. Setelah melalui proses *preprocessing* dan *filtering*, jumlah data yang semula sebanyak 2000 komentar berkurang menjadi 1087 komentar yang layak digunakan dalam proses analisis.

Selanjutnya dilakukan proses *tokenizing*, yaitu memisahkan teks komentar menjadi unit-unit kata (*token*). Setelah itu, teknik *stopword removal* diterapkan untuk menghapus kata-kata umum yang memiliki kontribusi relatif rendah terhadap analisis sentimen, seperti kata "yang", "dan", maupun "di". Tahap berikutnya adalah normalisasi kata, yaitu mengubah kata tidak baku atau kata slang ke dalam bentuk yang lebih sesuai dengan bahasa Indonesia, seperti "bgt" menjadi "banget", "gk" menjadi "tidak", dan "serem" menjadi "seram". Tahap terakhir adalah *stemming*, yaitu mengubah kata berimbuhan menjadi bentuk kata dasar sehingga kata-kata yang memiliki akar kata sama dapat direpresentasikan secara lebih konsisten.

Hasil tahapan pra-pemrosesan teks menunjukkan bahwa data komentar menjadi lebih bersih dan terorganisasi, sehingga siap untuk tahap ekstraksi fitur menggunakan pendekatan TF-IDF. Prosedur ini diharapkan dapat meningkatkan kualitas representasi data, yang pada gilirannya mempermudah model *Multinomial Naïve Bayes* dalam mengidentifikasi pola sentimen. Tabel 4.2 menampilkan contoh perubahan yang dilakukan pada komentar, baik sebelum maupun sesudah prosedur pra-pemrosesan teks.

Tabel 4. 2 Contoh Hasil Preprocessing

No	Komentar Asli	Hasil *Preprocessing*
1	Serem banget ya Allah SWT 😭😭😭	serem banget ya allah swt
2	Serem banget aku Ampe tutup mata sama telinga	serem banget aku ampe tutup mata sama telinga
3	Seru banget	seru banget

Berdasarkan Tabel 4.2, terlihat bahwa proses *text preprocessing* berhasil membersihkan data komentar dari karakter yang tidak diperlukan, seperti emoji, simbol, serta perbedaan penggunaan huruf kapital. Selain itu, proses *cleaning* juga mempertahankan kata-kata yang memiliki makna penting sehingga informasi utama pada komentar tetap terjaga. Hasil preprocessing ini menghasilkan data teks yang lebih bersih, konsisten, dan siap digunakan pada tahap ekstraksi fitur menggunakan metode TF-IDF.

Proses text *preprocessing* diimplementasikan menggunakan bahasa pemrograman Python pada platform Google Colab. Gambar 4.2 menunjukkan implementasi proses cleaning pada data komentar.

```
# Membuat salinan dataset hasil scraping
df_filter = df_scraping.copy()

def clean_text(text):
    text = str(text).lower()
    text = re.sub(r'<.*?>', '', text)
    text = re.sub(r'http\S+|www\S+', '', text)
    text = re.sub(r'[^a-zA-Z\s]', '', text)
    text = re.sub(r'\s+', ' ', text).strip()
    return text

df_filter['cleaned'] = df_filter['Komentar'].apply(clean_text)

df_filter[['Komentar', 'cleaned']].head(30)
```

Gambar 4. 2 Proses Cleaning (Preprocessing)

Gambar 4.2 menunjukkan implementasi proses *cleaning* menggunakan Python pada Google Colab. Proses tersebut digunakan untuk membersihkan komentar dari unsur-unsur yang tidak diperlukan sebelum data dilanjutkan ke tahapan *preprocessing* berikutnya.

4.1.3 Hasil Pelabelan Sentimen

Setelah proses *text preprocessing* selesai dilakukan, tahap berikutnya adalah pelabelan sentimen (*sentiment labeling*). Pelabelan dilakukan secara manual dengan mengelompokkan setiap komentar ke dalam tiga kategori, yaitu sentimen positif, netral, dan negatif. Hasil pelabelan ini digunakan sebagai dasar dalam proses klasifikasi menggunakan algoritma *Multinomial Naïve Bayes*. Beberapa contoh komentar pada masing-masing kategori sentimen ditampilkan pada gambar 4.3. di bawah ini

	Komentar	final_text	label
0	Seru banget	seru banget	Positif
1	Bagus nya film	bagus nya film	Positif
2	Aku sedih banget lihat adi nangis😢😢	aku sedih banget lihat adi nangis	Positif
3	setan opo iki kok mateni wong2	setan opo iki kok mateni wong	Positif
4	ngeri ngeri sedap	ngeri ngeri sedap	Positif
6	Banyak pelajaran yg diambil..	banyak pelajaran yang diambil	Positif
9	film ini pas tadi hari pahlawan aku nobar ini ...	film pas tadi hari pahlawan aku nobar selesai	Positif
10	Sumpah Cok baru pertama kali gua punya fans di...	sumpah cok baru pertama kali gua punya fans fi...	Positif
17	Kaya beneran	kaya beneran	Positif
18	Film ya bagus banget	film bagus banget	Positif

Gambar 4. 3 Contoh Komentar Sentimen Positif

Berdasarkan Gambar 4.3, terlihat beberapa contoh komentar yang telah diberi label sentimen positif. Komentar-komentar tersebut umumnya berisi ungkapan apresiasi, kepuasan, dan kesan baik terhadap film Waktu Maghrib, seperti “Seru banget”, “Bagus nya film”, dan “Film ya bagus banget”. Hasil pelabelan ini menunjukkan bahwa komentar yang mengandung opini positif digunakan sebagai data acuan (*ground truth*) dalam proses pelatihan dan pengujian model klasifikasi sentimen.

	Komentar	final_text	label
5	Masih bingung sma alurnya kok tiba² di teror 😕	bingung sama alurnya kok tiba teror	Negatif
7	Ngegantung bet anjr,nasib karta abis nyelamati...	ngegantung bet anjrnasib karta abis nyelamatin...	Negatif
13	Cerita ini cuman boongan percaya sangat 😂😂😂 ak...	cerita cuman boongan percaya sangat aku meliha...	Negatif
14	Ceritanya pabaliut	ceritanya pabaliut	Negatif
15	filem apaan coba siang kok; ada hantu 😂	filem apaan coba siang kok hantu	Negatif
21	Mass nya si Salman kayak anyinggg,kesell bet g...	mass nya si salman kayak anyingggkesell bet gw...	Negatif
22	Aduh sayang endingnya ini gk jelas	aduh sayang endingnya tidak jelas	Negatif
27	alur crita ga da, tjuan film ga ada, sutradara...	alur crita tidak da tjuan film tidak sutradara...	Negatif
28	Alur critanya gak jelas gak ada asal usul han...	alur critanya tidak jelas tidak asal usul hant...	Negatif
33	ini itu filmnnya gak beneran 😅😅	filmnnya tidak beneran	Negatif

Gambar 4. 4 Contoh Komentar Sentimen Negatif

Berdasarkan Gambar 4.4, komentar yang termasuk ke dalam sentimen negatif umumnya berisi kritik, kekecewaan, dan ketidakpuasan terhadap film. Sebagai contoh, komentar "Masih bingung sma alurnya kok tiba di teror" menunjukkan penilaian bahwa alur cerita sulit dipahami. Komentar "Aduh sayang endingnya ini gk jelas" mengungkapkan kekecewaan terhadap akhir cerita, sedangkan komentar "alur crita ga da, tjuan film ga ada, sutradara…" menunjukkan penilaian negatif terhadap alur, tujuan cerita, dan penyutradaraan film. Setelah melalui tahap preprocessing, komentar tersebut diubah menjadi teks yang lebih bersih sehingga dapat digunakan pada proses analisis sentiment

	Komentar	final_text	label
8	Itu palsu atau enggak😢	palsu enggak	Netral
11	ada penampakan	penampakan	Netral
12	Siapa yang masih nonton	siapa nonton	Netral
16	Memakai bahasa Jawa loh😲 aku bisa memakai baha...	memakai bahasa jawa loh aku memakai bahasa jawa	Netral
19	Jadi teringat	jadi teringat	Netral
25	Bujang inam	bujang inam	Netral
31	Akusukamu❤️😻😘😚🤩😊□□	akusukamu	Netral
36	Ali pacar aku	ali pacar aku	Netral
37	Ucapan adalah doa❤️	ucapan doa	Netral
40	terus nenek nya salman gimana tu ga ada yg ngu...	terus nenek nya salman gimana tu tidak yang ng...	Netral

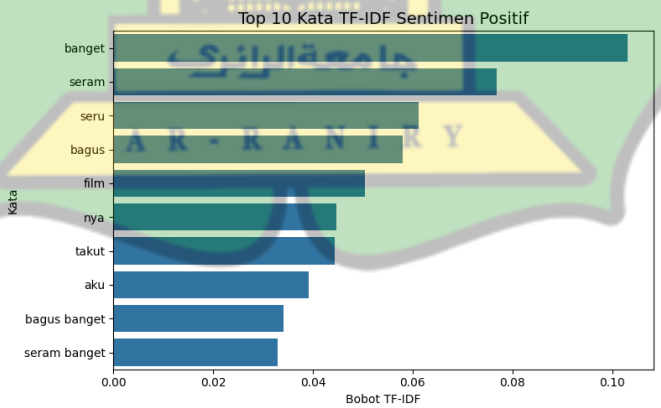
Gambar 4. 5 Contoh Komentar Sentimen Netral

Berdasarkan Gambar 4.5, komentar yang termasuk ke dalam sentimen netral umumnya tidak menunjukkan penilaian positif maupun

negatif terhadap film. Sebagai contoh, komentar “Siapa yang masih nonton” hanya berupa pertanyaan, komentar “Jadi teringat” sekadar mengungkapkan kesan tanpa memberikan evaluasi, dan komentar “Seram” merupakan respons singkat yang tidak secara jelas menunjukkan kepuasan ataupun ketidakpuasan terhadap film. Setelah melalui tahap preprocessing, komentar tersebut diubah menjadi teks yang lebih bersih sehingga dapat digunakan dalam proses analisis sentimen.

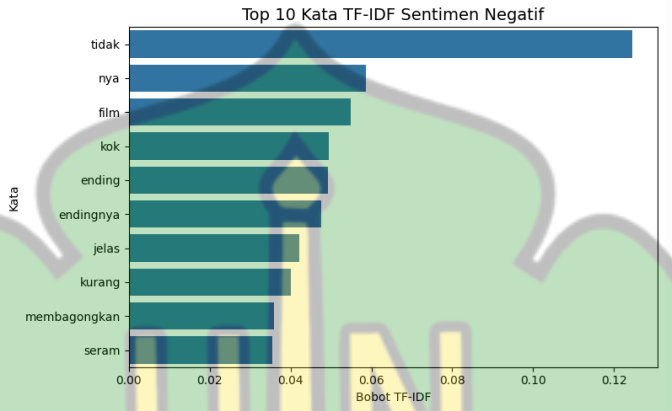
Berdasarkan Gambar 4.3 sampai Gambar 4.5, komentar yang telah melalui proses pelabelan berhasil dikelompokkan ke dalam tiga kategori sentimen, yaitu positif, netral, dan negatif. Pengelompokan tersebut dilakukan berdasarkan makna atau opini yang terkandung pada setiap komentar sehingga dapat digunakan sebagai data berlabel pada proses pelatihan dan pengujian model klasifikasi.

Setelah tahap pelabelan dan pra-pemrosesan, frekuensi kata dalam setiap kelas sentimen dianalisis. Tujuan tahap ini adalah untuk mengidentifikasi istilah-istilah yang paling sering muncul dalam komentar yang diklasifikasikan sebagai netral, negatif, atau positif. Data tersebut disajikan dalam bentuk daftar "10 Kata Teratas" guna memberikan gambaran mengenai karakteristik istilah-istilah kunci pada setiap kategori.



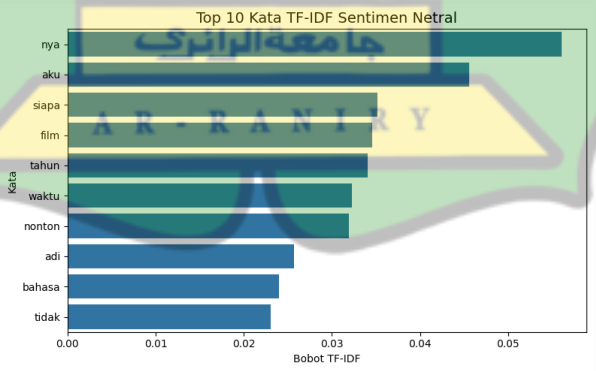
Gambar 4. 6 Top 10 Kata Sentimen Positif

Gambar 4.6 menunjukkan 10 kata dengan bobot TF-IDF tertinggi pada kelas sentimen positif. Kata “banget”, “seram”, dan “seru” menjadi tiga kata teratas yang paling mendominasi, yang mengindikasikan bahwa ulasan positif didominasi oleh ekspresi kepuasan dan antusiasme penonton terhadap film tersebut.



Gambar 4. 7 Top 10 Kata Sentimen Negatif

Gambar 4.7 menampilkan 10 kata dengan bobot TF-IDF tertinggi pada kategori sentimen negatif. Kata “tidak”, “nya”, dan “film” menempati posisi tiga teratas, yang mengindikasikan bahwa ulasan bernada negatif pada data tersebut didominasi oleh kritik atau keluhan yang berkaitan dengan kata-kata tersebut.



Gambar 4. 8 Top 10 Kata Sentimen Netral

Gambar 4.8 menyajikan 10 kata dengan bobot TF-IDF tertinggi pada kategori sentimen netral. Kata “nya”, “aku”, dan “siapa” menjadi kata yang paling sering muncul, menunjukkan bahwa ulasan bernada netral cenderung berupa pernyataan objektif, ajakan, atau cerita pengalaman biasa saat menonton film tanpa ekspresi emosional yang kuat.

4.1.4. Pembagian Data Training dan Testing

Setelah proses pelabelan sentimen selesai dilakukan, data hasil pelabelan manual kemudian dibagi menjadi dua bagian, yaitu data t*raining* dan data *testing*. Pembagian data dilakukan sebelum proses ekstraksi fitur menggunakan metode TF-IDF agar tidak terjadi *data leakage*, sehingga informasi dari data uji tidak memengaruhi proses pembelajaran model.

Pembagian data dilakukan dengan rasio 80:20, yaitu 80% data digunakan sebagai data *training* dan 20% sebagai data *testing*. Dengan menggunakan parameter *test_size* = 0.2 dan *random_state* = 42 pada fungsi *train_test_split* dari pustaka *Scikit-learn*, diperoleh 400 data *training* dan 100 data *testing* dari total 500 komentar berlabel. Parameter *stratify=y* digunakan untuk mempertahankan proporsi masing-masing kelas sentimen pada data *training* dan *testing*.

Penggunaan *random_state* = 42 bertujuan untuk menetapkan nilai awal (*seed*) pada proses pengacakan data sehingga pembagian dataset menghasilkan susunan data yang konsisten setiap kali program dijalankan dengan dataset yang sama. Angka 42 bukan merupakan nilai khusus atau ketentuan dalam metode *train-test split*, sehingga secara teknis dapat diganti dengan angka lain, seperti 1, 10, 100, atau nilai lainnya. Namun, penggunaan angka 42 dipertahankan dalam penelitian ini sebagai nilai *seed* yang telah ditetapkan untuk memastikan hasil pembagian data dapat direproduksi dan memudahkan proses pengujian ulang penelitian.

Selanjutnya, data *training* digunakan dalam proses pembentukan fitur TF-IDF dan pelatihan model *Multinomial Naïve Bayes*, sedangkan data *testing* digunakan untuk menguji kemampuan model dalam mengklasifikasikan data yang belum digunakan selama proses pelatihan.

Tabel 4.3 menyajikan pembagian data *training* dan *testing*.

Tabel 4. 3 Pembagian Data

Jenis Data	Jumlah
Data Training	400
Data Testing	100

4.1.5. Hasil Ekstraksi Fitur *(TF-IDF)*

Tahap selanjutnya adalah ekstraksi fitur menggunakan metode *Term Frequency–Inverse Document Frequency* (TF-IDF) setelah proses *text preprocessing*, pelabelan, dan pembagian data selesai dilakukan. Tahap ini bertujuan mengubah data teks menjadi representasi numerik sehingga dapat digunakan sebagai masukan dalam proses klasifikasi menggunakan algoritma *Multinomial Naïve Bayes*. Pembobotan TF-IDF dilakukan berdasarkan tingkat kemunculan suatu kata dalam dokumen (*Term Frequency*) dan tingkat kemunculan kata tersebut pada keseluruhan dokumen (*Inverse Document Frequency*).

Pada penelitian ini, ekstraksi fitur dilakukan menggunakan *TfidfVectorizer* dari pustaka *Scikit-learn* dengan parameter `ngram_range=(1,2)` dan `min_df=2`. Parameter `ngram_range=(1,2)` digunakan untuk mempertimbangkan unigram dan bigram, yaitu fitur yang terdiri dari satu kata dan gabungan dua kata yang berurutan. Penggunaan unigram dan bigram memungkinkan model memperoleh informasi tidak hanya dari kata secara individual, tetapi juga dari kombinasi dua kata yang dapat memberikan konteks tambahan dalam proses klasifikasi. Sementara itu, parameter `min_df=2` digunakan untuk mempertahankan fitur yang muncul pada sekurang-kurangnya dua dokumen sehingga kata atau kombinasi kata yang hanya muncul satu kali dapat dihilangkan.

Proses ekstraksi fitur dilakukan dengan menerapkan operasi *fit_transform* hanya pada data *training*, sedangkan data *testing* diproses menggunakan operasi *transform*. Dengan demikian, *vocabulary* dan bobot TF-IDF dibentuk berdasarkan data *training* dan tidak menggunakan informasi dari data *testing*. Penerapan proses tersebut bertujuan untuk mencegah terjadinya *data leakage* sehingga data *testing* tetap menjadi data yang belum diketahui oleh model selama proses pelatihan.

Berdasarkan hasil ekstraksi fitur, diperoleh matriks TF-IDF berukuran (400, 410) pada data *training* dan (100, 410) pada data *testing*. Hasil tersebut menunjukkan bahwa 400 data training dan 100 data testing masing-masing direpresentasikan menggunakan 410 fitur TF-IDF. Jumlah fitur pada data *training* dan *testing* sama karena data *testing* ditransformasikan menggunakan *vocabulary* yang telah dibentuk dari data *training*.

Representasi numerik hasil TF-IDF selanjutnya digunakan sebagai masukan dalam proses pelatihan dan pengujian algoritma *Multinomial Naïve Bayes*. Dengan demikian, setiap komentar telah diubah menjadi bentuk vektor numerik yang dapat diproses oleh model klasifikasi.

Tabel 4. 4 Informasi TF-IDF

Keterangan	Nilai
Data training	400
Data testing	100
Jumlah fitur TF-IDF	410

Berdasarkan Tabel 4.4, hasil ekstraksi fitur menggunakan metode TF-IDF menunjukkan bahwa data *training* sebanyak 400 komentar dan data *testing* sebanyak 100 komentar berhasil direpresentasikan menggunakan 410 fitur. Setiap komentar diubah menjadi bentuk vektor numerik berdasarkan bobot TF-IDF yang merepresentasikan tingkat kepentingan masing-masing fitur dalam dokumen. Representasi numerik tersebut selanjutnya digunakan

sebagai masukan dalam proses pelatihan algoritma *Multinomial Naïve Bayes*. Data *training* digunakan untuk membentuk dan melatih model, sedangkan data *testing* digunakan untuk menguji kemampuan model dalam mengklasifikasikan data yang belum digunakan selama proses pelatihan. Gambar 4.9 menampilkan hasil transformasi TF-IDF tersebut.

```
Ukuran data training: (400, 410)
Ukuran data testing: (100, 410)
```

Gambar 4. 9 Hasil Transformasi TF-IDF

Hasil transformasi data menggunakan pendekatan *Term Frequency–Inverse Document Frequency* (TF-IDF) ditampilkan pada Gambar 4.9. Berdasarkan hasil transformasi tersebut, dihasilkan matriks fitur dengan dimensi (400, 410) untuk data *training* dan (100, 410) untuk data *testing*. Hal ini menunjukkan bahwa 410 fitur berhasil direpresentasikan dari 400 sampel *training* dan 100 sampel *testing*. Fitur-fitur tersebut selanjutnya digunakan sebagai masukan dalam proses pelatihan dan pengujian model *Multinomial Naïve Bayes*.

4.1.6. Hasil Klasifikasi Menggunakan *Multinomial Naïve Bayes*

Setelah proses ekstraksi fitur menggunakan metode TF-IDF selesai dilakukan, tahap berikutnya adalah pelatihan model menggunakan algoritma *Multinomial Naïve Bayes*. Model dilatih menggunakan 400 data ***training*** yang telah direpresentasikan dalam bentuk TF-IDF dan selanjutnya digunakan untuk melakukan prediksi terhadap 100 data ***testing***. Hasil prediksi kemudian dibandingkan dengan label sebenarnya (*ground truth*) untuk mengetahui kemampuan model dalam mengklasifikasikan komentar ke dalam tiga kategori sentimen, yaitu positif, netral, dan negatif. Kinerja model selanjutnya dievaluasi menggunakan nilai *accuracy*, *precision*, *recall*, *F1-score*, serta *confusion matrix*. Hasil nilai akurasi model *Multinomial Naïve Bayes* ditampilkan pada Gambar 4.10.

```
Akurasi: 0.63
Akurasi (%): 63.0 %
```

Gambar 4. 10 Hasil Akurasi Multinomial Naïve Bayes

Berdasarkan Gambar 4.10, diperoleh nilai akurasi model *Multinomial Naïve Bayes* sebesar 0,63 atau 63%. Hasil tersebut menunjukkan bahwa model mampu mengklasifikasikan komentar ke dalam tiga kategori sentimen, yaitu positif, netral, dan negatif, dengan tingkat ketepatan sebesar 63% pada data *testing*. Nilai akurasi tersebut diperoleh dari hasil prediksi terhadap 100 data *testing* yang kemudian dibandingkan dengan label sebenarnya (*ground truth*).

Setelah dilakukan pengujian menggunakan model *Multinomial Naïve Bayes*, selanjutnya dilakukan evaluasi untuk mengetahui kinerja model dalam mengklasifikasikan sentimen komentar. Evaluasi dilakukan menggunakan beberapa metrik, yaitu *accuracy*, *precision*, *recall*, dan *F1-score*. Selain itu, *confusion matrix* digunakan untuk melihat jumlah prediksi yang benar dan salah pada masing-masing kelas sentimen. Hasil evaluasi diperoleh dengan membandingkan hasil prediksi model terhadap label sebenarnya (*ground truth*) pada data *testing*. Tabel 4.5 menyajikan hasil evaluasi model *Multinomial Naïve Bayes*.

Tabel 4. 5 Hasil Evaluasi Multinomial Naïve Bayes

Metrik	Nilai
Accuracy	63%
Precision (Weighted Avg)	65%
Recall (Weighted Avg)	63%
F1-Score (Weighted Avg)	60%

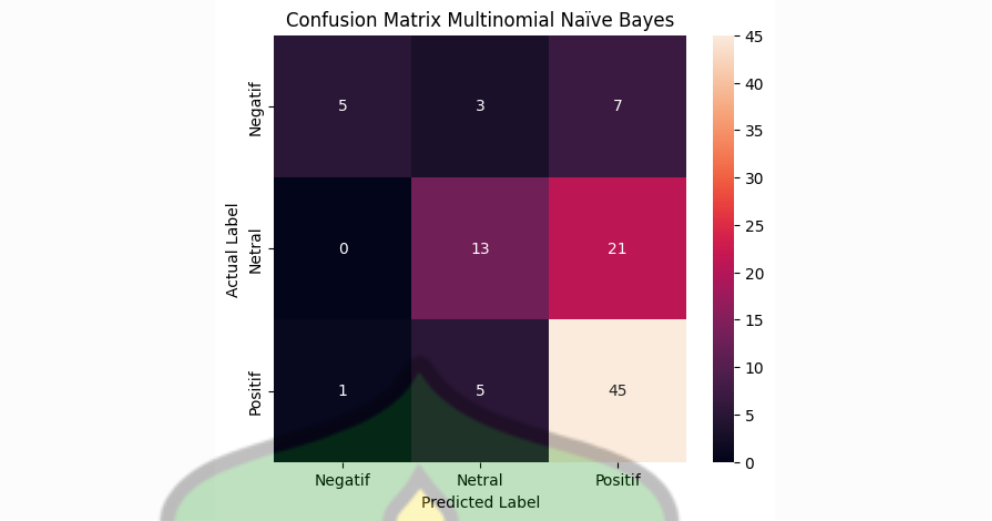
Berdasarkan Tabel 4.5, model *Multinomial Naïve Bayes* yang dilatih menggunakan 400 data *training* dan diuji menggunakan 100 data *testing* memperoleh nilai *accuracy* sebesar 63%. Selain itu, model menghasilkan

nilai *precision (weighted average)* sebesar 65%, *recall* sebesar 63%, dan *F1-score (weighted average)* sebesar 60%. Hasil tersebut menunjukkan bahwa model mampu mengklasifikasikan komentar ke dalam tiga kategori sentimen, yaitu positif, netral, dan negatif, dengan tingkat ketepatan sebesar 63% pada data *testing*. Perbedaan nilai *precision*, *recall*, dan *F1-score* pada setiap kelas menunjukkan bahwa kemampuan model dalam mengenali masing-masing kategori sentimen masih berbeda. Hasil evaluasi tersebut selanjutnya menjadi dasar untuk menganalisis kemampuan *Multinomial Naïve Bayes* dalam mengklasifikasikan sentimen berdasarkan karakteristik komentar yang digunakan dalam penelitian.

Untuk memperoleh gambaran performa model secara lebih rinci, hasil klasifikasi menggunakan *Multinomial Naïve Bayes* selanjutnya dievaluasi menggunakan *classification report* dan *confusion matrix*. *Classification report* menyajikan nilai *precision*, *recall*, dan *F1-score* pada setiap kelas sentimen, sedangkan *confusion matrix* menunjukkan jumlah prediksi yang benar maupun yang salah pada masing-masing kelas. Hasil evaluasi tersebut disajikan pada Gambar 4.11 dan Gambar 4.12.

```
Akurasi: 0.63
Akurasi (%): 63.0 %

Classification Report:
              precision    recall  f1-score   support

     Negatif       0.83      0.33      0.48        15
      Netral       0.62      0.38      0.47        34
     Positif       0.62      0.88      0.73        51

    accuracy                           0.63       100
   macro avg       0.69      0.53      0.56       100
weighted avg       0.65      0.63      0.60       100
```

Gambar 4. 11 Classification Report Model Multinomial Naïve Bayes



Gambar 4. 12 Confusion Matrix Model Multinomial Naïve Bayes

Hasil klasifikasi selanjutnya dianalisis menggunakan confusion matrix yang ditampilkan pada Gambar 4.12. *Confusion matrix* tersebut memberikan informasi mengenai jumlah data yang berhasil diprediksi dengan benar maupun yang mengalami kesalahan klasifikasi pada setiap kelas sentimen. Berdasarkan sajian data pada Gambar, ketepatan prediksi model mencakup 5 data berlabel negatif, 13 data netral, dan 45 data positif. Adapun bentuk deviasi prediksi mencakup kelas negatif aktual yang salah diprediksi sebagai netral (3 data) dan positif (7 data). Pada kategori netral aktual, deviasi terjadi ke arah positif sebanyak 21 data, dan tidak ada data netral yang salah diprediksi sebagai negatif (0 data). Sedangkan pada kategori positif aktual, kesalahan klasifikasi terdistribusi ke dalam kelas negatif (1 data) dan kelas netral (5 data).

Ketidakmampuan model dalam mengenali karakteristik setiap kelas secara tepat menunjukkan bahwa masih terdapat kesalahan klasifikasi pada beberapa kategori sentimen. Kesalahan tersebut terutama terlihat pada kelas netral yang masih cukup banyak diprediksi sebagai positif. Kondisi ini menunjukkan bahwa beberapa komentar netral memiliki karakteristik kata yang mirip dengan komentar positif sehingga model mengalami kesulitan dalam membedakan kedua kelas tersebut. Selain itu, masih terdapat kesalahan klasifikasi pada kelas negatif dan positif, yang menunjukkan bahwa model

belum sepenuhnya mampu mengenali karakteristik spesifik dari setiap kategori sentimen. Hal tersebut dapat dipengaruhi oleh variasi penggunaan bahasa, kata tidak baku, singkatan, serta konteks komentar yang digunakan dalam data penelitian.

4.2. Pembahasan

4.2.1. Pembahasan Hasil Klasifikasi *Multinomial Naïve Bayes*

Pada penelitian ini, proses klasifikasi sentimen dilakukan menggunakan algoritma *Multinomial Naïve Bayes* terhadap komentar penonton film *Waktu Maghrib* di YouTube. Model dilatih menggunakan data *training* yang telah melalui tahap *text preprocessing*, pelabelan sentimen, pembagian data, dan ekstraksi fitur TF-IDF. Selanjutnya, model diuji menggunakan data *testing* untuk mengetahui kemampuannya dalam mengklasifikasikan komentar ke dalam tiga kategori sentimen, yaitu positif, netral, dan negatif. Hasil klasifikasi kemudian dievaluasi menggunakan *accuracy*, *precision*, *recall*, *F1-score*, *classification report*, dan *confusion matrix* untuk mengetahui kinerja model secara keseluruhan maupun pada masing-masing kelas sentimen.

Berdasarkan hasil pengujian, model *Multinomial Naïve Bayes* memperoleh nilai *accuracy* sebesar 63%. Hasil tersebut menunjukkan bahwa model mampu mengklasifikasikan sebagian data pengujian ke dalam kategori sentimen yang sesuai dengan label sebenarnya. Meskipun demikian, masih terdapat kesalahan klasifikasi pada beberapa data, terutama karena adanya kemiripan karakteristik kata antar kelas sentimen. Oleh karena itu, hasil klasifikasi perlu dianalisis lebih lanjut berdasarkan nilai *precision*, *recall*, *F1-score*, *classification report*, dan *confusion matrix* untuk mengetahui kemampuan model dalam mengenali masing-masing kategori sentimen.

Karakteristik algoritma *Multinomial Naïve Bayes* yang bekerja berdasarkan distribusi probabilitas kata turut memengaruhi hasil klasifikasi. Model mengandalkan pola kemunculan kata pada data pelatihan untuk

menentukan kelas sentimen pada data pengujian. Oleh karena itu, ketika terdapat kata atau pola bahasa yang memiliki karakteristik serupa pada beberapa kelas, model dapat mengalami kesulitan dalam menentukan kategori sentimen yang tepat.

Selain itu, komentar pada media sosial memiliki karakteristik yang beragam, seperti penggunaan bahasa tidak baku, singkatan, kata gaul, serta kalimat yang mengandung makna ambigu. Kondisi tersebut menyebabkan beberapa komentar memiliki karakteristik yang mirip antar kelas sentimen sehingga masih sulit dibedakan oleh model. Hal ini juga terlihat pada hasil *classification report* dan *confusion matrix*, di mana masih terdapat sejumlah komentar yang diprediksi ke kelas sentimen yang tidak sesuai.

Secara keseluruhan, hasil penelitian menunjukkan bahwa *Multinomial Naïve Bayes* dapat digunakan untuk mengklasifikasikan sentimen komentar penonton film *Waktu Maghrib* ke dalam tiga kategori, yaitu positif, netral, dan negatif. Model memperoleh *accuracy* sebesar 63%, dengan *precision* (*weighted average*) sebesar 65%, *recall* sebesar 63%, dan *F1-score* sebesar 60%. Hasil tersebut menunjukkan bahwa model memiliki kemampuan dalam melakukan klasifikasi, meskipun masih terdapat kesalahan dalam membedakan beberapa komentar antar kelas sentimen.

4.2.2. Faktor yang Mempengaruhi Hasil Klasifikasi

Hasil klasifikasi sentimen pada penelitian ini dipengaruhi oleh beberapa faktor yang berkaitan dengan karakteristik data komentar dan metode yang digunakan. Faktor-faktor tersebut memengaruhi kemampuan model *Multinomial Naïve Bayes* dalam mengenali pola sentimen sehingga berdampak pada hasil klasifikasi yang diperoleh.

Salah satu faktor yang memengaruhi hasil klasifikasi adalah penggunaan bahasa tidak baku pada komentar YouTube. Pengguna sering menuliskan komentar dengan ejaan yang tidak sesuai dengan kaidah bahasa Indonesia, seperti penyingkatan kata, penggunaan huruf secara berulang,

maupun variasi penulisan tertentu. Meskipun telah dilakukan tahap *text preprocessing*, masih terdapat variasi kata yang dapat memengaruhi representasi fitur sehingga model mengalami kesulitan dalam mengenali pola sentimen secara tepat.

Selain itu, penggunaan singkatan dan kata slang juga menjadi faktor yang memengaruhi hasil klasifikasi. Komentar pada media sosial umumnya menggunakan istilah seperti *gk*, *ga*, *bgt*, *udh*, dan *bangettt*. Variasi penulisan tersebut menyebabkan kata yang memiliki makna serupa dapat direpresentasikan sebagai fitur yang berbeda. Kondisi ini dapat memengaruhi proses pembelajaran model dalam mengenali karakteristik setiap kelas sentimen.

Faktor berikutnya adalah adanya komentar yang bersifat ambigu. Beberapa komentar dapat mengandung lebih dari satu kecenderungan sentimen dalam satu kalimat, misalnya memadukan pujian dan kritik secara bersamaan. Kondisi tersebut menyebabkan model mengalami kesulitan dalam menentukan kelas sentimen yang paling sesuai sehingga masih ditemukan kesalahan klasifikasi pada data *testing*.

Karakteristik komentar pada film bergenre horor juga turut memengaruhi hasil klasifikasi. Kata-kata seperti "seram", "takut", dan "ngeri" tidak selalu menunjukkan sentimen negatif. Dalam konteks film horor, kata-kata tersebut dapat digunakan sebagai bentuk apresiasi karena menunjukkan bahwa film mampu memberikan suasana yang menegangkan. Namun, kata yang sama juga dapat digunakan dalam komentar yang mengandung kritik atau ketidakpuasan. Perbedaan konteks tersebut menjadi salah satu faktor yang menyebabkan model *Multinomial Naïve Bayes* mengalami kesulitan dalam membedakan sentimen hanya berdasarkan pola kemunculan kata.

Selain karakteristik bahasa, jumlah data berlabel yang digunakan dalam penelitian juga dapat memengaruhi performa model. Dari 1.087 komentar hasil *preprocessing* dan *filtering*, sebanyak 500 komentar dipilih

sebagai sampel untuk dilakukan pelabelan manual (*ground truth*). Selanjutnya, data tersebut dibagi menjadi 400 data *training* dan 100 data *testing*. Jumlah data berlabel tersebut masih relatif terbatas untuk merepresentasikan seluruh variasi bahasa dan pola sentimen yang terdapat pada komentar YouTube. Kondisi ini dapat membatasi kemampuan model dalam mempelajari karakteristik setiap kelas sentimen.

Secara keseluruhan, hasil penelitian menunjukkan bahwa performa *Multinomial Naïve Bayes* dipengaruhi oleh penggunaan bahasa tidak baku, singkatan dan kata slang, komentar ambigu, karakteristik istilah pada film horor, serta jumlah data berlabel yang digunakan. Faktor-faktor tersebut dapat menyebabkan adanya kemiripan karakteristik antar kelas sehingga model masih mengalami kesalahan dalam mengklasifikasikan beberapa komentar. Meskipun demikian, model mampu melakukan klasifikasi ke dalam tiga kategori sentimen dengan nilai *accuracy* sebesar 63%, sehingga dapat digunakan untuk menganalisis sentimen komentar penonton film *Waktu Maghrib*.

4.2.3. Kelebihan dan Keterbatasan Model *Multinomial Naïve Bayes*

Berdasarkan hasil penelitian yang telah dilakukan, algoritma *Multinomial Naïve Bayes* memiliki beberapa kelebihan dalam proses klasifikasi sentimen komentar penonton film *Waktu Maghrib* di YouTube. Salah satu kelebihannya adalah proses pelatihan dan pengujian model yang relatif sederhana, cepat, serta efisien dalam mengolah data teks. Selain itu, algoritma ini mampu bekerja dengan baik pada data berdimensi tinggi yang telah direpresentasikan menggunakan metode TF-IDF, sehingga sesuai untuk diterapkan pada penelitian analisis sentimen.

Model Multinomial Naïve Bayes juga mampu mengklasifikasikan komentar ke dalam tiga kategori sentimen, yaitu positif, netral, dan negatif. Pada pengujian model, diperoleh nilai accuracy sebesar 63%. Hasil tersebut menunjukkan bahwa model mampu melakukan proses klasifikasi sentimen

dengan cukup baik, meskipun performanya masih dipengaruhi oleh karakteristik data teks yang digunakan.

Di sisi lain, model ini juga memiliki beberapa keterbatasan. Salah satu kelemahan utama Multinomial Naïve Bayes adalah asumsi bahwa setiap fitur atau kata bersifat independen. Dalam kenyataannya, makna suatu komentar sering kali dipengaruhi oleh hubungan antar kata dan konteks kalimat. Akibatnya, model masih mengalami kesulitan dalam mengklasifikasikan komentar yang mengandung makna ambigu, penggunaan bahasa tidak baku, singkatan, maupun kata-kata slang yang umum digunakan pada media sosial. Selain itu, pada komentar film bergenre horor, kata-kata spesifik seperti “seram” atau “takut” sering kali dapat bermakna ganda (bisa berupa apresiasi positif maupun kritik negatif) yang sulit dibedakan oleh model hanya berdasarkan frekuensi kemunculan kata.

Secara keseluruhan, Multinomial Naïve Bayes merupakan algoritma yang sederhana dan efektif untuk analisis sentimen berbasis teks. Namun, berdasarkan hasil penelitian ini, performanya masih dapat ditingkatkan melalui penggunaan dataset berlabel yang lebih banyak, proses normalisasi bahasa dan penanganan kata slang yang lebih baik, serta penerapan algoritma klasifikasi lain yang mampu menangani konteks dan karakteristik data teks secara lebih optimal.

4.2.4. Implikasi Hasil Penelitian

Hasil penelitian ini menunjukkan bahwa algoritma Multinomial Naïve Bayes dapat diterapkan untuk melakukan analisis sentimen terhadap komentar penonton film Waktu Maghrib pada platform YouTube. Melalui tahapan text preprocessing, ekstraksi fitur menggunakan metode TF-IDF, serta proses klasifikasi, model mampu mengelompokkan komentar ke dalam tiga kategori sentimen, yaitu positif, netral, dan negatif dengan tingkat akurasi sebesar 63%. Hasil ini membuktikan bahwa Multinomial

Naïve Bayes merupakan metode yang sederhana dan efisien untuk digunakan dalam klasifikasi sentimen berbasis teks.

Dari sisi praktis, hasil penelitian ini dapat dimanfaatkan sebagai bahan evaluasi bagi rumah produksi film maupun pembuat konten untuk mengetahui kecenderungan opini penonton berdasarkan komentar yang diberikan di media sosial. Informasi mengenai proporsi sentimen positif, netral, dan negatif dapat membantu dalam memahami respons serta persepsi publik terhadap film Waktu Maghrib, sehingga dapat menjadi masukan berharga dalam pengembangan kualitas cerita, teknik penyutradaraan, maupun penyusunan strategi promosi pada karya berikutnya.

Dari sisi akademis, penelitian ini memberikan gambaran mengenai penerapan kombinasi TF-IDF dan Multinomial Naïve Bayes pada teks ulasan film bergenre horor di media sosial berbahasa Indonesia. Penelitian ini memperlihatkan bahwa meskipun algoritma berbasis frekuensi kata mampu memberikan gambaran umum opini penonton, karakteristik data teks berupa penggunaan bahasa tidak baku, kata slang, dan kalimat ambigu tetap menjadi tantangan tersendiri dalam proses klasifikasi automatik.

Meskipun model Multinomial Naïve Bayes mampu mengklasifikasikan data dengan cukup baik, masih terdapat keterbatasan yang memengaruhi performanya. Berdasarkan evaluasi confusion matrix, masih ditemukan kesalahan prediksi pada komentar yang mengandung konteks kompleks, makna ambigu, maupun kata-kata horor spesifik yang dapat dipersepsikan secara positif maupun negatif. Kondisi tersebut menunjukkan bahwa model masih mengalami kesulitan dalam membedakan komentar yang memiliki kemiripan makna antar kelas sentimen secara presisi.

Oleh karena itu, penelitian ini dapat menjadi dasar bagi pengembangan sistem analisis sentimen pada komentar YouTube, khususnya pada ulasan film. Pengembangan selanjutnya dapat dilakukan melalui normalisasi bahasa yang lebih optimal, penambahan jumlah data

berlabel, serta penerapan metode penyeimbangan data atau algoritma klasifikasi lain yang lebih sesuai dengan karakteristik dataset. Dengan demikian, diharapkan performa klasifikasi sentimen pada penelitian selanjutnya dapat menjadi lebih akurat dan konsisten.

BAB V
PENUTUP

5.1 Kesimpulan

Penelitian ini menyimpulkan bahwasanya pendekatan Multinomial Naïve Bayes efektif digunakan untuk mengelompokkan sentimen komentar audiens film Waktu Maghrib pada platform YouTube menjadi tiga kategori, yaitu sentimen positif, netral, dan negatif. Proses penelitian dilakukan secara terstruktur melalui tahapan pengumpulan data menggunakan YouTube Data API, text preprocessing, pelabelan sentimen secara manual, pembagian data, ekstraksi fitur menggunakan TF-IDF, proses klasifikasi, hingga evaluasi model. Hasil penelitian menunjukkan bahwa model Multinomial Naïve Bayes mampu melakukan klasifikasi sentimen dengan baik berdasarkan pola kata dan pembobotan TF-IDF yang terdapat pada komentar pengguna.

Berdasarkan hasil evaluasi model, analisis sentimen menggunakan algoritma Multinomial Naïve Bayes dan pembobotan fitur TF-IDF memperoleh tingkat akurasi sebesar 63%. Hasil tersebut menunjukkan bahwa model mampu mengenali dan memprediksi sentimen komentar penonton film Waktu Maghrib dengan cukup baik. Kesalahan klasifikasi yang terjadi pada beberapa data uji umumnya disebabkan oleh adanya kata-kata ambigu, penggunaan bahasa tidak baku/slang, serta konteks kalimat khas film horor yang sulit dibedakan oleh model hanya berdasarkan frekuensi dan pembobotan kata. Secara keseluruhan, Multinomial Naïve Bayes membuktikan keandalannya sebagai algoritma yang sederhana, cepat, dan efisien dalam melakukan klasifikasi sentimen berbasis teks.

5.2 Saran

Berdasarkan kesimpulan dan keterbatasan dalam penelitian ini, saran yang dapat diberikan untuk penerapan maupun pengembangan penelitian selanjutnya adalah sebagai berikut:

1. Bagi Mahasiswa atau Praktisi

Hasil penelitian ini dapat dimanfaatkan sebagai salah satu referensi dalam memahami kecenderungan sentimen penonton terhadap film melalui komentar di platform YouTube. Bagi rumah produksi film maupun pihak yang berkepentingan, hasil analisis sentimen dapat digunakan sebagai bahan evaluasi untuk mengetahui respons penonton terhadap film yang telah dipublikasikan sehingga dapat menjadi masukan dalam meningkatkan kualitas maupun strategi promosi film di masa mendatang.

2. Bagi Peneliti Selanjutnya

Pengembangan penelitian mendatang direkomendasikan untuk memperluas cakupan sampel data (dataset) dengan menyertakan beragam unggahan video serta mengekstraksi opini dari berbagai jejaring media sosial platform lain, sehingga generalisasi dan tingkat representativitas temuan dapat ditingkatkan. Selain itu, penelitian berikutnya dapat membandingkan metode *Multinomial Naïve Bayes* dengan algoritma klasifikasi lain, seperti *Support Vector Machine (SVM)*, *Random Forest*, atau *Long Short-Term Memory (LSTM)* untuk mengetahui metode yang memberikan performa terbaik dalam analisis sentimen.

3. Bagi Pengembangan Penelitian

Penelitian selanjutnya juga disarankan untuk menerapkan teknik penyeimbangan data selain *Random Oversampling*, seperti *SMOTE (Synthetic Minority Over-sampling Technique)* atau metode lainnya, serta mengembangkan tahapan *preprocessing* dengan menambahkan normalisasi bahasa gaul, singkatan, emoji, dan kata tidak baku. Penggunaan data yang lebih beragam dan proses pengolahan teks yang lebih optimal diharapkan dapat meningkatkan akurasi model serta menghasilkan performa klasifikasi sentimen yang lebih baik.

DAFTAR PUSTAKA

Alwan, D., & Ridla, M. A. (2024). Jurnal Sistem dan Teknologi Informasi Indonesia Averaged Word2vec sebagai Ekstraksi Fitur pada Analisis Sentimen Ulasan Film di IMDb menggunakan Artificial Neural Network (ANN) Averaged Word2vec as Feature Extraction in Sentiment Analysis of Movie Revie. 9(1),36–45. https://doi.org/10.32528/justindo.v9i1.1204

Y Ayu Putri Gabriella S. (2023). *Optimasi Penerimaan Siswa Baru dengan Penerapan Algortima Text Mining dan TF-IDF*. *2*(3), 109–115. https://doi.org/10.47065/comforch.v2i3.941

Astuti, K. C. (2024). Implementasi Text Mining Untuk Analisis Sentimen Masyarakat Terhadap Ulasan Aplikasi Digital Korlantas Polri pada Google Play Store. REMIK: Riset Dan E-Jurnal Manajemen Informatika Komputer, 8(1), 383–394.https://www.jurnal.polgan.ac.id/index.php/remik/article/view/13421

Bintoro, P., Ratnasari, Wihardjo, E., Putri, I. P., & Asari, A. (2024). *Pengantar machine learning*. PT Mafy Media Literasi Indonesia. https://repository.um.ac.id/id/eprint/5619

Firsttama, R. A., Arifiyanti, A. A., & Kartika, D. S. Y. (2024). Analisis sentimen komentar YouTube Konferensi Tingkat Tinggi G20 menggunakan metode Naive Bayes. *Jurnal Teknologi dan Sistem Informasi Bisnis*, *6*(2), 282–285. https://doi.org/10.47233/jteksis.v6i2.1263

Fitrianti, S., & Yudhistira, A. (2025). Analisis sentimen media sosial terhadap calon Pilkada 2024 dengan metode Naïve Bayes. Jurnal Pendidikan dan Teknologi Indonesia, 5(1), 167–176. https://doi.org/10.52436/1.jpti.610

Hidayatullah, & Jadiaman Parhusip. (2024). Analisis Distribusi Selisih Tingkat Pengganggguran Terbuka (Tpt) Dan Tingat Partisipasi Angkatan Kerja (Tpak) Menggunakan Google Colab. Informatika: Jurnal Teknik Informatika Dan Multimedia, 4(2), 39–45. https://doi.org/10.51903/informatika.v4i2.828

Hermawan, A., & Jowensen, J. (2023). Text preprocessing dalam text mining: Teknik pembersihan dan persiapan data untuk analisis teks. Jurnal Teknologi Informasi dan Ilmu Komputer, 10(3), 210–218. https://doi.org/10.23887/jstundiksha.v12i1.52358

Iedwan, A. S., Mauliza, N., Pristyanto, Y., Hartanto, A. D., & Rohman, A. N. (2024). Comparative performance of SVM and Multinomial Naïve Bayes in sentiment analysis of the film *Dirty Vote*. *Scientific Journal of Informatics, 11*(3), 839–848. https://doi.org/10.15294/sji.v11i3.10290

Nurpandi, F., Sulaeman, F. S., & Hermawan, A. (2024). *Analisis sentimen terhadap kinerja kepolisian Indonesia menggunakan metode Multinomial Naive Bayes, Long Short-Term Memory, dan Lexicon-Based*. Media Jurnal Informatika, 16(1). http://jurnal.unsur.ac.id/mjinformatika

Nitha Kumala Dewi. (2023). Identifikasi Berita Hoax dengan Menerapkan Algoritma Text Mining. Journal of Informatics, Electrical and Electronics Engineering, 2(3), 65–74. https://doi.org/10.47065/jieee.v2i3.888

Nur Cahyo, D., Zulfia Zahro', H., & Vendyansyah, N. (2023). Pengenalan Ekspresi Mikro Wajah Dengan Ekstraksi Fitur Pada Komponen Wajah Menggunakan Metode Local Binary Pattern Histogram. JATI (Jurnal Mahasiswa Teknik Informatika), 7(1), 822–829. https://doi.org/10.36040/jati.v7i1.6167

Palupi, L., Ihsanto, E., & Nugroho, F. (2023). Analisis Validasi dan Evaluasi

Model Deteksi Objek Varian Jahe Menggunakan Algoritma Yolov5. Journal of Information System Research (JOSH), 5(1), 234–241. https://doi.org/10.47065/josh.v5i1.4380

Pratama, D. R. S., Munandar, T. A., & Ramdhania, K. F. (2024). *Multinomial Naive Bayes algorithm for Indonesian language sentiment classification related to Jakarta International Stadium (JIS)*. *International Journal of Information Technology and Computer Science Applications, 2*(1), 12–22. https://doi.org/10.58776/ijitcsa.v2i1.118

Pambudi, R. E., Purnomo, H., & Irawan, R. (2024). Pemanfaatan Data Mining Untuk Prediksi Prestasi Akademik Siswa Berdasarkan Pola Kehadiran , Aktivitas Belajar. 16(2), 132–141. https://doi.org/10.32767/jti.v16i2.2507

Pebdika, A., Herdiana, R., & Solihudin, D. (2023). Klasifikasi Menggunakan Metode Naive Bayes Untuk Menentukan Calon Penerima Pip. JATI (Jurnal Mahasiswa Teknik Informatika), 7(1), 452–458. https://doi.org/10.36040/jati.v7i1.6303

Prasetiyowati, S. S., & Sibaroni, Y. (2024). Unlocking the potential of Naive Bayes for spatio temporal classification: A novel approach to feature expansion. Journal of Big Data, 11, 106. https://doi.org/10.1186/s40537-024-00958-x

Salman Al Markas, M., Pratama, R., & Hidayat, A. (2025). Penerapan metode Bag of Words dalam representasi teks untuk analisis sentimen. Jurnal Informatika dan Sistem Informasi, 6(1), 45–52. https://doi.org/10.33096/linier.v2i2.3104

Saragih, N. N., & Kurniawan, R. (2025). *Analisis sentimen menggunakan metode Naïve Bayes tentang program mudik gratis Pemerintah Kota Medan 2024. CESS (Journal of Computer Engineering, System and Science), 10*(1), 299–311. https://doi.org/10.24114/cess.v10i1.65930

Salim, E., & Syafrullah, M. (2023). Analisis sentimen pada ulasan pelayanan Suku Dinas Kependudukan dan Pencatatan Sipil Kota Administrasi Jakarta Barat menggunakan algoritme K-Nearest Neighbor. *Bit (Fakultas Teknologi Informasi Universitas Budi Luhur), 20*(1), 58–65. https://doi.org/10.36080/bit.v20i1.2186

Suhendra, A., Rahman, F., & Putri, D. (2024). Analisis sentimen teks berbahasa Indonesia menggunakan metode Multinomial Naïve Bayes. *Jurnal Teknologi Informasi dan Ilmu Komputer, 11*(2), 145–152. https://journal.nurulfikri.ac.id/index.php/JIT/article/download/1544/383

Wati, R., Ernawati, S., & Rachmi, H. (2023). Pembobotan TF-IDF Menggunakan Naïve Bayes pada Sentimen Masyarakat Mengenai Isu Kenaikan BIPIH. Jurnal Manajemen Informatika (JAMIKA), 13(1), 84–93. https://doi.org/10.34010/jamika.v13i1.9424

Wibisono, A. D. R., Mandyartha, E. P., & Al Haromainy, M. M. (2025). Klasifikasi Penyakit Kulit Berbasis Support Vector Machine Dengan Ekstraksi Fitur Abcd Rule. JIPI (Jurnal Ilmiah Penelitian Dan Pembelajaran Informatika), 10(1), 686–698. https://doi.org/10.29100/jipi.v10i1.6039

Wahyudi, M. F. D., Pratama, A., & Saputra, R. (2024). Pemanfaatan Google Colab dalam analisis data berbasis cloud computing. Jurnal Teknologi Informasi dan Komunikasi, 8(2), 120–128. https://doi.org/10.51903/informatika.v4i2.828

LAMPIRAN

Lampiran 1 Source Code Pengambilan Data (Scraping) Menggunakan YouTube Data API

- **Cell Import Library**

```
import pandas as pd
import numpy as np
import re

from googleapiclient.discovery import build

from langdetect import detect, DetectorFactory
DetectorFactory.seed = 0

import nltk
nltk.download('punkt')

from nltk.tokenize import word_tokenize
from Sastrawi.StopWordRemover.StopWordRemoverFactory import
StopWordRemoverFactory
from Sastrawi.Stemmer.StemmerFactory import StemmerFactory

from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.model_selection import train_test_split
from sklearn.naive_bayes import MultinomialNB
from sklearn.metrics import classification_report,
confusion_matrix, accuracy_score

from imblearn.over_sampling import RandomOverSampler

import matplotlib.pyplot as plt
import seaborn as sns
```

- **Kode Pengambilan Komentar (Scraping)**

```
api_key = input("Masukkan API Key: ")
youtube = build('youtube', 'v3', developerKey=api_key)
video_id = "uFGE0hMw2OQ"
```

- **Kode Penyimpanan ke CSV**

```
df.to_csv('komentar_mentah.csv', index=False)
print("Berhasil! File komentar_mentah.csv sudah tersimpan di
folder Colab.")
```

- **Membaca hasil scraping**

```
df_scraping = pd.read_csv("komentar_mentah.csv")
```

Lampiran 2 Source Code Text Preprocessing

- **Cleaning Text**

```
df_filter = df_scraping.copy()
def clean_text(text):
text = str(text).lower()
text = re.sub(r'<.*?>', '', text)
text = re.sub(r'http\S+|www\S+', '', text)
text = re.sub(r'[^a-zA-Z\s]', '', text)
text = re.sub(r'\s+', ' ', text).strip()
return text
df_filter['cleaned'] = df_filter['Komentar'].apply(clean_text)
df_filter[['Komentar', 'cleaned']].head(30)
```

- **Filtering Bahasa**

```
def detect_lang(text):
    try:
        return detect(text)
    except:
        return "unknown"

df_filter['lang'] = df_filter['cleaned'].apply(detect_lang)

df_filter = df_filter[df_filter['lang'] == 'id']

print("Total komentar setelah filtering bahasa:",
len(df_filter))
```

- **Tokenizing**

```
df_filter['tokenized'] =
df_filter['cleaned'].apply(word_tokenize)
```

- **Stopword Removal**

```
from Sastrawi.StopWordRemover.StopWordRemoverFactory import
StopWordRemoverFactory

factory = StopWordRemoverFactory()
stopwords = set(factory.get_stop_words())
```

- **Stemming**

```
stem_factory = StemmerFactory()
stemmer = stem_factory.create_stemmer()
```

- **Pengambilan Sampel 500 Data**

```
df_sample = df_filter.sample(n=500, random_state=42)

df_sample = df_sample[['Komentar', 'final_text']].copy()

df_sample['label'] = ""
```

Lampiran 3 Source Code Pemodelan

- **Train-Test Split**

```
from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(
    X,
    y,
    test_size=0.2,
    random_state=42,
    stratify=y
)

print("Jumlah data train:", len(X_train))
print("Jumlah data test:", len(X_test))
```

- **TF-IDF**

```
from sklearn.feature_extraction.text import TfidfVectorizer

tfidf = TfidfVectorizer()

X_train = tfidf.fit_transform(X_train)

X_test = tfidf.transform(X_test)

print("Shape X_train TF-IDF:", X_train.shape)
print("Shape X_test TF-IDF:", X_test.shape)
```

- **Multinomial Naïve Bayes**

```
nb = MultinomialNB()

nb.fit(X_train, y_train)

y_pred = nb.predict(X_test)

print("Akurasi Baseline:", accuracy_score(y_test, y_pred))
print(classification_report(y_test, y_pred))
```

- **Random Oversampling**

```
from imblearn.over_sampling import RandomOverSampler

ros = RandomOverSampler(random_state=42)
```

```
X_train_res, y_train_res = ros.fit_resample(X_train, y_train)
```

- **Multinomial Naïve Bayes Setelah Random Oversampling**

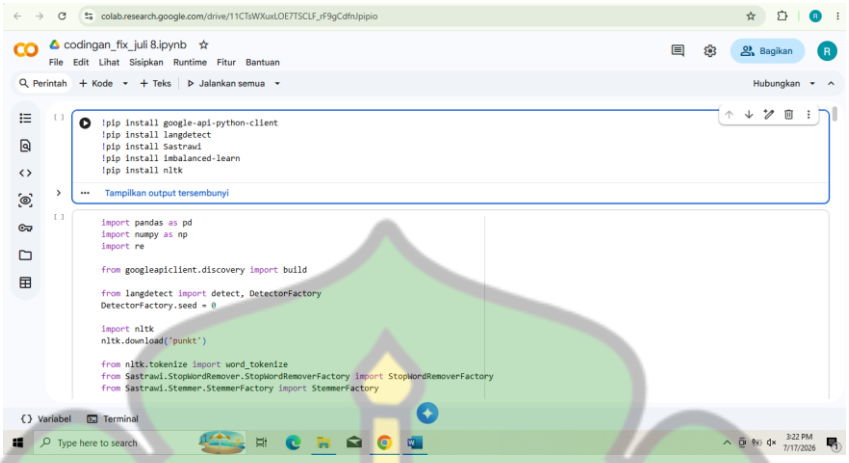
```
nb_ros = MultinomialNB()

nb_ros.fit(X_train_res, y_train_res)

y_pred_ros = nb_ros.predict(X_test)

print("Akurasi:", accuracy_score(y_test, y_pred_ros))
print(classification_report(y_test, y_pred_ros))
```

- **Evaluasi Model (Confusion Matrix)**

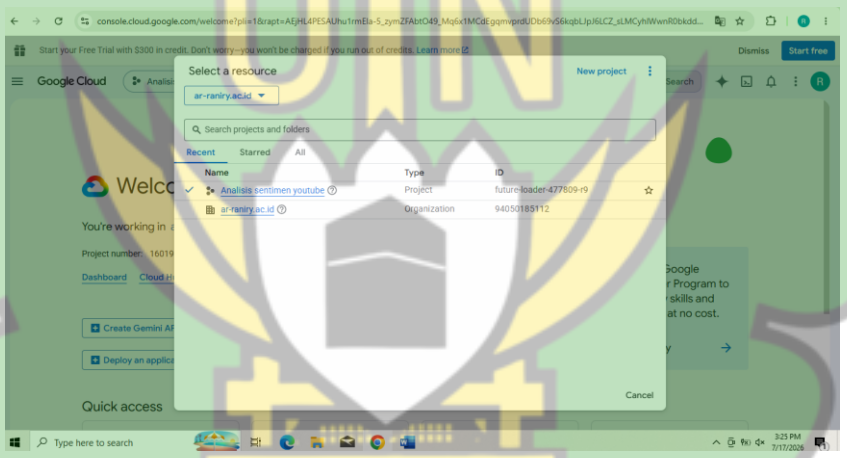
```
cm = confusion_matrix(y_test, y_pred_ros)

labels = ['Negatif', 'Netral', 'Positif']

plt.figure(figsize=(6,4))

sns.heatmap(
    cm,
    annot=True,
    fmt='d',
    xticklabels=labels,
    yticklabels=labels
)

plt.title("Confusion Matrix Multinomial Naive Bayes + Random
Oversampling")
plt.xlabel("Prediksi Model")
plt.ylabel("Label Sebenarnya")

plt.show()
```

Lampiran 4 Dokumentasi Penelitian

- **Halaman Google Colab saat menjalankan program**



- **Halaman Google Cloud Console (YouTube Data API aktif)**



- **Screenshot video YouTube Waktu Maghrib**