Indonesian Journal on Software Engineering (IJSE)
Vol. 11, No. 1, Juni 2025, hlm. 44-53 p-ISSN: 2461-0690
Akreditasi KEMENRISTEKDIKTI, No. 225/E/KPT/2022 e-ISSN: 2714-9935

Evaluasi Sentimen Ulasan CGV Cinemas Indonesia dengan Naïve Bayes dan SVM

Muhamad Ali Zaenal Abidin[1], Astriana Mulyani[2]

[1,2]Program Studi Informatika, Fakultas Teknologi Informasi, Universitas Nusa Mandiri
Email: [1]12240137@nusamandiri.ac.id, [2]astriana.atm@nusamandiri.ac.id

Abstrak

Perkembangan teknologi digital telah mendorong masyarakat untuk menggunakan aplikasi pemesanan tiket bioskop secara daring, termasuk CGV Cinemas Indonesia. Namun, meningkatnya jumlah ulasan dari pengguna aplikasi di Google Play Store memunculkan tantangan dalam mengidentifikasi persepsi pengguna secara efisien. Penelitian ini bertujuan untuk menganalisis sentimen pengguna terhadap aplikasi CGV Cinemas Indonesia dan membandingkan performa algoritma Naïve Bayes dan Support Vector Machine (SVM) dalam klasifikasi sentimen. Data dikumpulkan dari Google Play Store sebanyak 4.600 ulasan (2.300 ulasan positif dan 2.300 negatif) melalui teknik web scraping. Proses analisis mencakup tahap preprocessing teks, balancing, ekstraksi fitur menggunakan TF-IDF, pelatihan model dengan Python dan Google Colab, serta evaluasi kinerja model menggunakan metrik akurasi, presisi, recall, dan F1-score. Hasil penelitian menunjukkan bahwa Naïve Bayes lebih unggul dibandingkan SVM, dengan akurasi sebesar 90,38% dan F1-score sebesar 90,46%, sedangkan SVM memperoleh akurasi sebesar 88,88% dan F1-score 89,08%. Kesimpulan dari penelitian ini adalah bahwa metode Naïve Bayes lebih efektif dalam klasifikasi sentimen ulasan pengguna aplikasi CGV. Penelitian ini memberikan kontribusi dalam penerapan analisis sentimen untuk mendukung evaluasi layanan digital berbasis opini pengguna secara otomatis dan akurat.

Kata kunci: *Analisis Sentimen, Naïve Bayes, Support Vector Machine, CGV Cinemas Indonesia, Google Play Store*

Abstract

The advancement of digital technology has encouraged the public to use online movie ticket booking applications, including CGV Cinemas Indonesia. However, the increasing number of user reviews on the Google Play Store poses a challenge in efficiently identifying user perceptions. This study aims to analyze user sentiment towards the CGV Cinemas Indonesia application and compare the performance of Naïve Bayes and Support Vector Machine (SVM) algorithms in sentiment classification. The data were collected from Google Play Store, consisting of 4,600 reviews (2,300 positive and 2,300 negative), using web scraping techniques. The analysis process includes text preprocessing, data balancing, feature extraction using TF-IDF, model training with Python and Google Colab, and model evaluation using accuracy, precision, recall, and F1-score metrics. The results show that Naïve Bayes outperforms SVM, achieving an accuracy of 90.38% and an F1-score of 90.46%, while SVM obtained an accuracy of 88.88% and an F1-score of 89.08%. This research concludes that Naïve Bayes is more effective in classifying user sentiment on CGV app reviews. This study contributes to the application of sentiment analysis in supporting the evaluation of digital services based on user-generated opinions in an automated and accurate manner.

Keywords: *Sentiment Analysis, Naïve Bayes, Support Vector Machine, CGV Cinemas Indonesia, Google Play Store*

1. PENDAHULUAN

Perkembangan teknologi informasi dan komunikasi yang pesat pada era digital saat ini telah membawa dampak signifikan terhadap berbagai aspek kehidupan manusia, termasuk dalam cara masyarakat mengakses dan menikmati layanan hiburan. Salah satu bentuk transformasi digital yang paling

nyata dapat ditemukan pada industri perfilman, khususnya dalam hal pemesanan tiket bioskop yang kini telah beralih dari sistem konvensional menjadi sistem daring (online). Jika pada masa lalu masyarakat harus datang langsung ke lokasi bioskop dan mengantre untuk mendapatkan tiket, kini proses tersebut dapat dilakukan hanya dalam beberapa langkah mudah melalui perangkat digital seperti ponsel pintar (smartphone) atau tablet. Proses ini menjadi lebih mudah dengan adanya aplikasi mobile yang tersedia di berbagai platform digital seperti Google Play Store untuk pengguna Android dan App Store untuk pengguna iOS.

Salah satu aplikasi pemesanan tiket bioskop yang telah cukup lama beroperasi di Indonesia adalah aplikasi CGV Cinemas Indonesia. Aplikasi ini merupakan bagian dari layanan digital yang disediakan oleh jaringan bioskop CGV, yang bertujuan untuk memberikan kemudahan dan efisiensi kepada para penonton dalam melakukan berbagai aktivitas yang berhubungan dengan layanan bioskop. Melalui aplikasi ini, pengguna dapat melihat jadwal tayang film, memilih jenis studio dan kursi yang diinginkan, serta melakukan transaksi pembelian tiket secara langsung tanpa perlu hadir di lokasi. Fitur-fitur yang ditawarkan juga meliputi notifikasi film terbaru, promo menarik, dan integrasi dengan metode pembayaran digital. Kemudahan-kemudahan ini menjadi daya tarik tersendiri, khususnya bagi generasi muda yang terbiasa menggunakan layanan digital dalam aktivitas sehari-hari (Yusuf Rismanda Gaja et al., 2024).

Seiring dengan meningkatnya jumlah pengguna yang mengunduh dan menggunakan aplikasi tersebut, jumlah ulasan atau review yang diberikan oleh pengguna di halaman aplikasi juga mengalami peningkatan yang signifikan. Ulasan-ulasan ini mencerminkan berbagai opini dan pengalaman pengguna terkait aspek-aspek penting dari aplikasi, mulai dari tampilan antarmuka (user interface), kecepatan layanan, stabilitas aplikasi, hingga kualitas layanan pelanggan. Banyak dari ulasan tersebut yang menyampaikan pendapat jujur dan aktual mengenai kekuatan maupun kelemahan aplikasi, menjadikannya sebagai sumber data yang sangat berharga bagi pengembang untuk melakukan evaluasi dan peningkatan kualitas (Pebdika et al., 2023).

Namun, meskipun keberadaan ulasan pengguna ini sangat bermanfaat, data tersebut memiliki sifat yang tidak terstruktur dan volumenya yang terus bertambah setiap harinya menjadikannya sulit untuk dianalisis secara manual. Untuk membaca dan mengevaluasi ribuan komentar satu per satu secara manusiawi tentu akan membutuhkan waktu dan tenaga yang tidak sedikit. Selain itu, proses manual ini juga sangat rentan terhadap subjektivitas penilai yang bisa berbeda-beda dalam memahami makna atau emosi yang terkandung dalam suatu ulasan (Eldo et al., 2024). Oleh karena itu, dibutuhkan pendekatan yang lebih efisien, sistematis, dan objektif untuk mengekstrak informasi penting dari ulasan-ulasan tersebut, salah satunya adalah melalui teknik sentiment analysis.

Sentiment analysis atau analisis sentimen merupakan cabang dari bidang Natural Language Processing (NLP) yang bertujuan untuk mengidentifikasi, mengekstraksi, dan mengklasifikasikan opini atau emosi dari suatu teks. Dalam konteks ini, ulasan pengguna akan dipetakan ke dalam kategori tertentu seperti sentimen positif, negatif, atau netral, tergantung pada isi dari komentar tersebut (P. K. Sari & Suryono, 2024). Teknik ini telah banyak diterapkan di berbagai bidang, terutama dalam dunia bisnis dan pemasaran digital, karena mampu menyajikan gambaran umum tentang bagaimana persepsi masyarakat terhadap suatu produk atau layanan (Astuti et al., 2024).

Agar proses analisis sentimen dapat berjalan secara optimal, pemilihan metode klasifikasi yang digunakan dalam pemrosesan data teks menjadi hal yang sangat krusial. Dua metode yang paling banyak digunakan dalam klasifikasi data teks adalah algoritma Naïve Bayes dan Support Vector Machine (SVM). Algoritma Naïve Bayes merupakan metode berbasis probabilistik yang bekerja dengan prinsip Teorema Bayes dan mengasumsikan independensi antar fitur. Meskipun asumsi ini cukup sederhana, Naïve Bayes dikenal memiliki performa yang cukup baik dalam klasifikasi teks, bahkan ketika jumlah data latih tidak terlalu besar atau data yang tersedia tidak seimbang (Asyhari et al., 2023), (Haq et al., 2024).

Di sisi lain, Support Vector Machine adalah metode pembelajaran terawasi (supervised learning) yang bekerja dengan mencari hyperplane optimal yang memisahkan data ke dalam dua kelas yang berbeda dengan margin terbesar. Keunggulan utama dari SVM adalah kemampuannya dalam menangani data berdimensi tinggi dan non-linear dengan menggunakan fungsi kernel, seperti linear, polynomial, dan radial basis function (RBF), yang membantu memetakan data ke ruang berdimensi yang lebih tinggi untuk meningkatkan akurasi klasifikasi (Apriliyani et al., 2024), (Gumilar et al., 2024).

Berbagai penelitian sebelumnya telah menunjukkan bahwa kedua metode ini memiliki performa yang saling bersaing dalam konteks sentiment analysis. Sebagai contoh, penelitian (Betesda et al., 2024) yang melakukan analisis terhadap ulasan pada website google play menunjukkan bahwa SVM memberikan hasil akurasi yang lebih tinggi dibandingkan dengan Naïve Bayes, meskipun Naïve Bayes unggul dalam kecepatan komputasi. Penelitian lain (Lubis & Setyawan, 2024) terhadap ulasan aplikasi Pospay menunjukkan bahwa Naïve Bayes memiliki nilai F1-score yang lebih tinggi, menandakan adanya

keseimbangan yang lebih baik antara presisi dan recall. Selain itu, penelitian (Java et al., 2024) menegaskan pentingnya tahap preprocessing seperti penghapusan slangwords, normalisasi kata, dan pemilihan fitur untuk meningkatkan kualitas hasil klasifikasi. Lebih lanjut, penelitian (Fazrian et al., 2024) terhadap aplikasi game mencatat bahwa Naïve Bayes mampu mencapai akurasi hingga 98,96%, yang merupakan indikasi kekuatan metode ini dalam menangani data teks berbahasa Indonesia.

Namun demikian, dari berbagai kajian literatur yang telah dilakukan, belum ditemukan penelitian yang secara spesifik membandingkan performa algoritma Naïve Bayes dan SVM dalam menganalisis sentimen ulasan pengguna aplikasi CGV Cinemas Indonesia yang diunduh melalui Google Play Store. Kekosongan ini menunjukkan adanya research gap yang layak untuk diteliti lebih lanjut. Dengan membandingkan kedua metode tersebut dalam konteks aplikasi bioskop CGV, penelitian ini diharapkan dapat memberikan kontribusi nyata dalam pengembangan metode evaluasi layanan digital berbasis opini pengguna di Indonesia.

Adapun tujuan dari penelitian ini adalah untuk melakukan analisis sentimen terhadap ulasan pengguna aplikasi CGV Cinemas Indonesia dengan menggunakan algoritma Naïve Bayes dan Support Vector Machine, serta membandingkan performa keduanya berdasarkan metrik evaluasi seperti akurasi, presisi, recall, dan F1-score. Harapan dari penelitian ini adalah memberikan informasi yang berguna bagi pengembang aplikasi dalam menentukan metode klasifikasi yang paling efektif dan efisien, serta memperkaya literatur ilmiah di bidang analisis sentimen berbahasa Indonesia.

2. METODE PENELITIAN

Dalam upaya untuk menganalisis sentimen pengguna terhadap aplikasi CGV Cinemas Indonesia, serta untuk membandingkan performa dua algoritma klasifikasi teks, penelitian ini dirancang menggunakan pendekatan yang sistematis dan terstruktur. Tahapan metodologi yang diterapkan mencakup identifikasi jenis penelitian, pengumpulan dan pengolahan data, penerapan algoritma, hingga evaluasi performa model menggunakan metrik yang telah ditentukan. Pemilihan metode dan pendekatan ini disesuaikan dengan tujuan utama penelitian, yaitu mengevaluasi secara kuantitatif kemampuan klasifikasi dari dua algoritma pembelajaran mesin yang berbeda terhadap data ulasan dalam Bahasa Indonesia. Seluruh proses dilaksanakan secara digital dengan bantuan perangkat lunak dan pustaka yang relevan, guna memastikan validitas, efisiensi, serta reprodusibilitas hasil (R. Y. Sari et al., 2024).

2.1. Jenis dan Pendekatan Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimen sebagai dasar untuk mengkaji dan membandingkan performa dua algoritma klasifikasi dalam analisis sentimen terhadap ulasan aplikasi CGV Cinemas Indonesia yang diambil dari platform Google Play Store. Pendekatan kuantitatif dipilih karena sesuai untuk menghasilkan data numerik yang dapat dianalisis secara statistik. Metode eksperimen digunakan untuk mengamati perubahan kinerja algoritma yang diuji berdasarkan berbagai parameter evaluasi. Dua algoritma yang dibandingkan dalam penelitian ini adalah Naïve Bayes dan Support Vector Machine (SVM), yang masing-masing telah dikenal luas dalam domain klasifikasi teks dan sentiment analysis. Penggunaan metode eksperimen memungkinkan peneliti untuk melakukan proses pelatihan dan pengujian model dengan kondisi terkontrol, sehingga hasil yang diperoleh dapat dibandingkan secara objektif dan terukur.

Seluruh proses eksperimen dilakukan secara daring dengan memanfaatkan platform Google Colaboratory, yang memungkinkan integrasi langsung dengan pustaka machine learning dalam bahasa pemrograman Python. Platform ini dipilih karena menyediakan lingkungan notebook yang interaktif dan mendukung penggunaan pustaka penting seperti Scikit-learn, Pandas, dan Sastrawi. Google Colab juga memberikan akses gratis ke sumber daya komputasi berbasis GPU, yang mempercepat proses pelatihan model secara signifikan. Dengan demikian, pendekatan ini diharapkan mampu memberikan hasil analisis yang efisien dan mendalam dalam konteks klasifikasi sentimen berbasis teks ulasan aplikasi.

2.2. Sumber dan Teknik Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang bersumber dari ulasan pengguna terhadap aplikasi CGV Cinemas Indonesia yang tersedia di halaman Google Play Store. Pemilihan sumber ini didasarkan pada pertimbangan bahwa ulasan pengguna di platform tersebut bersifat aktual, autentik, dan merepresentasikan pengalaman langsung dari konsumen terhadap fitur, performa, dan layanan aplikasi. Untuk mengumpulkan data tersebut, peneliti menggunakan teknik web scraping dengan memanfaatkan pustaka Python bernama google-play-scraper, yang mampu mengekstraksi informasi dari halaman Google Play secara otomatis dan sistematis.

Pengumpulan data dilakukan dengan fokus pada dua kategori sentimen, yaitu positif dan negatif. Untuk memastikan kualitas dan kejelasan data, peneliti mengambil ulasan yang memiliki rating tertinggi (bintang 5) sebagai representasi sentimen positif dan rating terendah (bintang 1) sebagai representasi sentimen negatif. Masing-masing kategori dikumpulkan sebanyak 2.300 data, sehingga total jumlah data yang diperoleh adalah 4.600 ulasan. Pemilihan ulasan berdasarkan rating ekstrem ini bertujuan untuk menghindari ambiguitas sentimen, karena ulasan dengan rating tengah (2, 3, atau 4) cenderung memiliki isi yang ambigu dan dapat menurunkan akurasi pelabelan otomatis.

2.3. Teknik Pra-Pemrosesan Data

Sebelum digunakan dalam proses pelatihan model, data ulasan yang dikumpulkan harus melalui tahapan pra-pemrosesan (preprocessing) agar menjadi lebih bersih, terstruktur, dan siap diolah oleh algoritma pembelajaran mesin. Proses ini penting karena teks ulasan sering kali mengandung berbagai elemen yang tidak relevan atau mengganggu, seperti simbol, angka, emotikon, dan variasi huruf kapital. Tahapan pertama adalah cleaning, yaitu menghapus karakter non-alfabet seperti angka, simbol, tanda baca, emoji, dan tautan, agar teks bersih dari elemen yang tidak memiliki makna semantik.

Selanjutnya, dilakukan case folding, yakni proses mengubah seluruh karakter dalam teks menjadi huruf kecil (lowercase) untuk menghindari redundansi dalam analisis kata. Setelah itu, proses normalisasi kata dilakukan dengan mengganti kata tidak baku atau kata slang menjadi padanan kata baku sesuai dengan kaidah Bahasa Indonesia. Langkah berikutnya adalah tokenisasi, yaitu memecah kalimat menjadi unit-unit kata (tokens) sehingga setiap kata dapat dianalisis secara individual.

Tahapan lanjutan adalah stopword removal, yaitu menghapus kata-kata umum yang tidak memberikan kontribusi makna signifikan terhadap sentimen, seperti "yang", "dan", "di", serta kata hubung lainnya. Proses ini membantu mengurangi dimensi data dan meningkatkan efisiensi model. Terakhir, dilakukan stemming menggunakan pustaka Sastrawi, yang bertugas mengembalikan setiap kata ke bentuk dasarnya. Sebagai contoh, kata "berjalan" akan dikembalikan menjadi "jalan". Dengan tahap preprocessing yang sistematis, teks ulasan dapat direpresentasikan dalam format yang optimal untuk kebutuhan klasifikasi.

2.4. Penyeimbangan Data (Balancing)

Salah satu tantangan utama dalam klasifikasi adalah distribusi kelas yang tidak seimbang, yang dapat menyebabkan model menjadi bias terhadap kelas yang dominan. Dalam penelitian ini, untuk memastikan distribusi kelas yang seimbang antara sentimen positif dan negatif, dilakukan proses penyeimbangan data (balancing) dengan metode undersampling. Dari total 2.300 data pada masing-masing kelas, dipilih secara acak sebanyak 2.000 data untuk setiap kelas, sehingga total data yang digunakan dalam pelatihan dan pengujian model adalah 4.000 data. Teknik ini digunakan agar model tidak condong hanya pada satu jenis sentimen saja dan dapat melakukan prediksi yang lebih adil dan akurat terhadap kedua kelas yang ada.

2.5. Ekstraksi Fitur dengan TF-IDF

Setelah tahap preprocessing selesai dan data seimbang, langkah selanjutnya adalah mengubah data teks menjadi representasi numerik yang dapat diproses oleh algoritma klasifikasi. Proses ini dilakukan melalui teknik ekstraksi fitur dengan menggunakan Term Frequency-Inverse Document Frequency (TF-IDF). TF-IDF digunakan untuk mengukur pentingnya suatu kata dalam dokumen tertentu relatif terhadap keseluruhan korpus. Kata yang sering muncul dalam satu dokumen tetapi jarang dalam dokumen lain akan memperoleh bobot yang lebih tinggi.

Konfigurasi TF-IDF dalam penelitian ini mencakup penggunaan unigram dan bigram, yang berarti satu dan dua kata berturut-turut akan dihitung sebagai fitur (ngram_range=(1,2)). Selain itu, untuk menghindari fitur yang terlalu sering muncul (umum) atau terlalu jarang (unik), digunakan parameter max_df=0.95 dan min_df=5, yang berarti hanya kata-kata yang muncul di bawah 95% dan lebih dari 5 dokumen yang dipertahankan sebagai fitur. Representasi TF-IDF ini menghasilkan matriks berdimensi tinggi, namun efektif dalam menangkap informasi semantik dari teks.

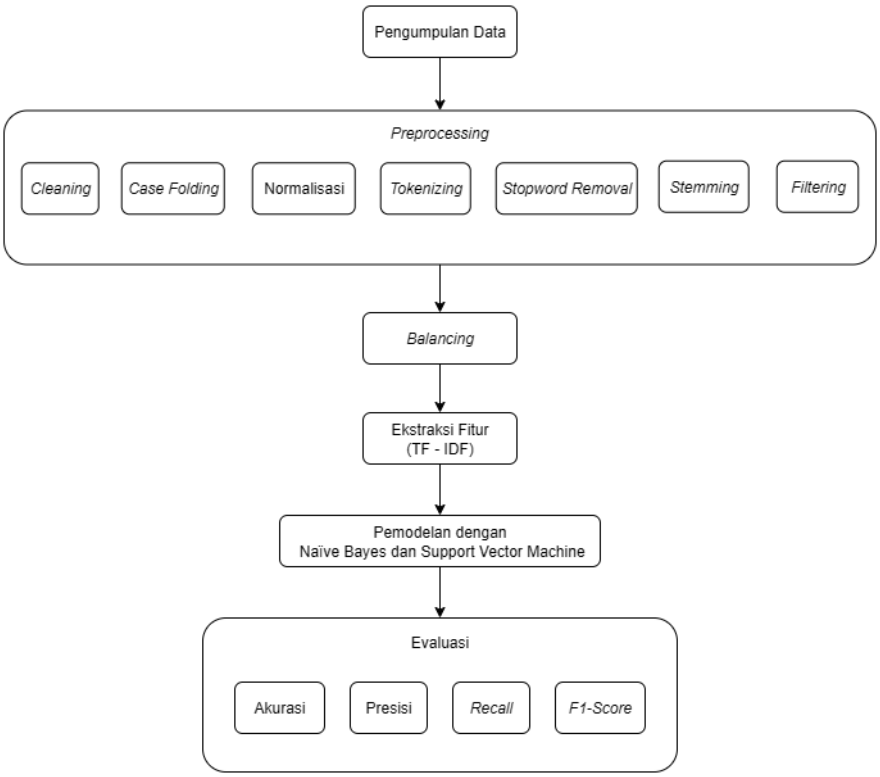
2.6. Pelatihan dan Pengujian Model

Model klasifikasi kemudian dibangun dengan menggunakan dua algoritma, yaitu Multinomial Naïve Bayes dan Linear Support Vector Classifier (LinearSVC). Sebelum pelatihan dilakukan, data dibagi menjadi dua bagian, yaitu 80% sebagai data latih dan 20% sebagai data uji. Pembagian ini bertujuan agar model dapat belajar dari mayoritas data dan diuji pada data yang belum pernah dilihat sebelumnya untuk mengukur kemampuan generalisasi.

Pelatihan model dilakukan pada lingkungan Google Colab dengan menggunakan pustaka Scikit-learn. Untuk mendapatkan hasil terbaik, dilakukan proses penyetelan parameter (hyperparameter tuning) dengan menggunakan GridSearchCV. Teknik ini memungkinkan pencarian kombinasi parameter yang optimal melalui proses evaluasi silang (cross-validation) sehingga performa model dapat dimaksimalkan(Maarif & Setiyawati, 2024).

2.7. Evaluasi Model

Setelah model selesai dilatih, kinerjanya dievaluasi dengan menggunakan empat metrik utama klasifikasi, yaitu akurasi, presisi, recall, dan F1-score. Akurasi mengukur seberapa banyak prediksi yang benar dibandingkan dengan seluruh data. Presisi mengukur ketepatan prediksi positif, yaitu seberapa banyak prediksi positif yang benar-benar positif. Recall mengukur seberapa banyak data positif yang berhasil dikenali dengan benar oleh model. Sementara itu, F1-score merupakan rata-rata harmonik dari presisi dan recall, yang sangat berguna ketika terdapat ketidakseimbangan antara jumlah kelas (Xu et al., 2024). Keempat metrik ini memberikan gambaran yang menyeluruh terhadap performa model dalam melakukan klasifikasi sentimen terhadap ulasan pengguna aplikasi CGV Cinemas Indonesia.



Gambar 1. Teknik Analisis Data

3. HASIL DAN PEMBAHASAN

Penelitian ini menyajikan hasil penerapan metode Naïve Bayes dan Support Vector Machine (SVM) dalam analisis sentimen terhadap ulasan pengguna aplikasi CGV Cinemas Indonesia. Seluruh tahapan dimulai dari pengumpulan data, pra-pemrosesan teks, ekstraksi fitur, pelatihan model, hingga evaluasi performa klasifikasi menggunakan metrik-metrik seperti akurasi, presisi, recall, dan F1-score. Bab ini juga menguraikan proses visualisasi hasil klasifikasi melalui confusion matrix, serta menyajikan analisis kesalahan yang terjadi selama proses prediksi. Pembahasan disusun untuk memberikan pemahaman yang mendalam mengenai efektivitas masing-masing algoritma dalam menangani data ulasan berbahasa Indonesia, serta mengidentifikasi faktor-faktor yang mempengaruhi keberhasilan atau kegagalan model dalam mengklasifikasikan opini pengguna secara otomatis.

3.1. Deskripsi Data Ulasan

Penelitian ini menggunakan data ulasan pengguna aplikasi CGV Cinemas Indonesia yang diperoleh dari platform Google Play Store. Proses pengambilan data dilakukan menggunakan teknik web scraping

dengan bantuan pustaka google-play-scraper dalam bahasa pemrograman Python. Pustaka ini memungkinkan peneliti untuk mengekstraksi ribuan ulasan secara otomatis dan efisien, lengkap dengan informasi seperti nama pengguna, tanggal, isi ulasan, dan rating yang diberikan. Setelah proses scraping, data yang berhasil dikumpulkan berjumlah 4.600 ulasan, terdiri dari 2.300 ulasan positif dengan rating bintang 5, dan 2.300 ulasan negatif dengan rating bintang 1. Ulasan dengan rating netral atau ambigu (bintang 2, 3, atau 4) tidak digunakan untuk menjaga kejelasan kategori sentimen dalam proses klasifikasi. Seluruh data ulasan disimpan dalam format Comma-Separated Values (CSV) agar mudah diolah dan diintegrasikan ke dalam pipeline analisis menggunakan Python. Penggunaan data yang cukup besar ini memberikan representasi yang lebih kuat terhadap persepsi pengguna terhadap aplikasi, sekaligus meningkatkan validitas model dalam mengidentifikasi pola sentimen.

3.2. Proses Pra-Pemrosesan Data

Sebelum digunakan dalam pelatihan dan pengujian model klasifikasi, data ulasan yang diperoleh dari proses scraping harus melalui tahapan pra-pemrosesan atau preprocessing. Tujuan utama dari tahap ini adalah untuk membersihkan dan menormalisasi data teks agar dapat digunakan secara optimal oleh algoritma pembelajaran mesin. Tahapan pertama adalah cleaning, yang bertujuan menghapus karakter-karakter non-alfabet, termasuk simbol-simbol, angka, emoji, URL, tanda baca, serta karakter khusus lainnya yang tidak relevan terhadap analisis sentimen. Misalnya, teks seperti “Bagus bangettt!!! 😊👍” akan dibersihkan menjadi “bagus banget”.

Langkah berikutnya adalah case folding, yaitu mengubah semua huruf dalam teks menjadi huruf kecil. Hal ini penting karena dalam Bahasa Indonesia huruf kapital dan huruf kecil memiliki makna yang sama, dan perbedaan kapitalisasi bisa menyebabkan fitur dianggap berbeda. Kemudian dilakukan normalisasi kata, yakni proses mengganti kata tidak baku atau slang menjadi kata baku sesuai Kamus Besar Bahasa Indonesia. Sebagai contoh, kata “nggak” diubah menjadi “tidak”, “bgt” menjadi “banget”, dan sebagainya.

Langkah selanjutnya adalah tokenisasi, yang memecah kalimat menjadi unit-unit kata terpisah. Setelah itu dilakukan stopword removal, yaitu menghapus kata-kata umum seperti “yang”, “dan”, “di”, yang tidak memiliki makna penting terhadap klasifikasi sentimen. Terakhir adalah stemming, yaitu mengubah setiap kata ke bentuk dasarnya dengan menggunakan pustaka Sastrawi. Proses ini misalnya mengubah kata “berjalan”, “menjalankan”, dan “dijalankan” menjadi “jalan”. Contoh nyata dari hasil pra-pemrosesan adalah: Ulasan awal “Aplikasinya jelek banget! Loading terus dan nggak bisa booking!” akan berubah menjadi “aplikasi jelek loading booking”. Proses ini sangat penting untuk memastikan bahwa hanya kata-kata bermakna yang digunakan dalam proses pembelajaran mesin, serta mengurangi redundansi dalam representasi fitur.

3.3. Penyeimbangan Data

Meskipun jumlah data awal yang dikumpulkan sudah seimbang antara sentimen positif dan negatif (masing-masing 2.300), pembagian data untuk pelatihan dan pengujian model tetap memerlukan proses penyesuaian. Data dibagi secara proporsional dengan rasio 80:20, di mana 80% digunakan untuk pelatihan dan 20% untuk pengujian. Dari total 4.600 data, sebanyak 3.680 data digunakan untuk melatih model (1.840 positif dan 1.840 negatif), sedangkan sisanya yaitu 920 data digunakan untuk menguji kinerja model.

Penyeimbangan ini dilakukan untuk memastikan bahwa tidak ada bias terhadap salah satu kelas sentimen, serta menjamin keadilan dalam proses evaluasi model. Model klasifikasi yang dilatih dengan data tidak seimbang cenderung memiliki performa yang buruk dalam mengidentifikasi kelas minoritas, yang menyebabkan turunnya nilai precision dan recall. Oleh karena itu, distribusi kelas yang proporsional merupakan syarat penting untuk menghasilkan model yang adil dan dapat diandalkan.

3.4. Ekstraksi Fitur dengan TF-IDF

Setelah tahap pra-pemrosesan selesai, data teks ulasan diubah menjadi representasi numerik agar dapat diproses oleh algoritma klasifikasi. Proses ini dilakukan melalui metode Term Frequency-Inverse Document Frequency (TF-IDF), yang merupakan teknik pembobotan kata berdasarkan frekuensi kemunculan dalam dokumen tertentu dibandingkan seluruh korpus. TF-IDF memberikan bobot tinggi pada kata-kata yang unik dan informatif, sementara mengabaikan kata-kata yang terlalu umum. Misalnya, kata “film” mungkin muncul di hampir semua ulasan, sehingga akan memiliki bobot rendah, sedangkan kata seperti “lag” atau “lemot” yang muncul hanya dalam ulasan negatif akan mendapat bobot tinggi.

Konfigurasi parameter yang digunakan adalah ngram_range=(1,2) untuk menangkap baik unigram (kata tunggal) maupun bigram (dua kata berurutan), max_df=0.95 untuk mengabaikan kata yang muncul terlalu sering, dan min_df=5 untuk hanya menyertakan kata-kata yang muncul minimal lima kali. Representasi TF-IDF ini kemudian menghasilkan matriks fitur berdimensi tinggi yang menggambarkan

keterkaitan antar kata dan dokumen, yang selanjutnya akan menjadi input utama dalam pelatihan model klasifikasi.

3.5. Pelatihan dan Evaluasi Model

Setelah data teks dikonversi menjadi vektor numerik, dilakukan proses pelatihan terhadap dua model klasifikasi, yaitu Naïve Bayes (MultinomialNB) dan Support Vector Machine (LinearSVC). Pelatihan dilakukan pada 3.680 data dan pengujian pada 920 data. Model pertama, Multinomial Naïve Bayes, adalah algoritma berbasis probabilistik yang menghitung peluang suatu dokumen termasuk dalam kelas tertentu berdasarkan distribusi kata. Model kedua, Linear Support Vector Classifier, bekerja dengan menemukan garis pemisah (hyperplane) terbaik yang memisahkan dua kelas dalam ruang vektor berdimensi tinggi.

Pelatihan dilakukan di lingkungan Google Colaboratory menggunakan pustaka scikit-learn. Awalnya, model dijalankan dengan parameter default untuk mendapatkan baseline performance. Selanjutnya, dilakukan penyetelan parameter (hyperparameter tuning) menggunakan metode GridSearchCV dengan validasi silang sebanyak 5-fold untuk mencari kombinasi parameter terbaik. Seluruh model dievaluasi menggunakan empat metrik utama yaitu accuracy, precision, recall, dan F1-score.

3.6. Hasil Evaluasi Naïve Bayes dan SVM

Hasil evaluasi menunjukkan bahwa kedua model menghasilkan performa yang tinggi, namun Naïve Bayes sedikit lebih unggul dalam seluruh metrik. Untuk model Naïve Bayes, diperoleh hasil sebagai berikut: akurasi sebesar 90,38%, presisi 90,61%, recall 90,31%, dan F1-score 90,46%. Angka ini menunjukkan bahwa Naïve Bayes mampu mengklasifikasikan ulasan pengguna secara konsisten dan akurat.

Sementara itu, model SVM menghasilkan akurasi sebesar 88,88%, presisi 89,04%, recall 88,77%, dan F1-score 89,08%. Meski performanya juga sangat baik, nilai-nilai ini sedikit lebih rendah dibandingkan dengan hasil model Naïve Bayes. Perbedaan ini dapat disebabkan oleh karakteristik teks ulasan yang relatif pendek dan memiliki pola bahasa yang sederhana, yang lebih sesuai dengan pendekatan probabilistik dari Naïve Bayes.

3.7. Perbandingan Performa Naïve Bayes dengan SVM

Secara umum, perbandingan hasil menunjukkan bahwa Naïve Bayes lebih unggul dalam mengklasifikasikan sentimen ulasan aplikasi CGV dibandingkan dengan SVM. Hal ini terlihat dari keunggulan konsisten pada metrik akurasi, presisi, recall, maupun F1-score. Salah satu faktor utama keunggulan Naïve Bayes adalah kemampuannya dalam menangani teks pendek dengan jumlah fitur terbatas, seperti yang umum ditemukan dalam ulasan aplikasi.

Di sisi lain, SVM dikenal memiliki performa optimal pada data berdimensi tinggi dan kompleksitas spasial yang lebih besar. Namun, dalam konteks ini, penggunaan unigram dan bigram menghasilkan fitur yang relatif sederhana, sehingga keunggulan SVM dalam hal margin optimal tidak terlalu tampak. Selain itu, waktu pelatihan Naïve Bayes juga jauh lebih cepat dibandingkan SVM, menjadikannya pilihan yang efisien untuk implementasi skala besar.

3.8. Pembahasan Implikasi Hasil

Hasil penelitian ini memiliki implikasi yang signifikan dalam tiga ranah utama, yaitu pengembangan aplikasi, penelitian lanjutan, dan kontribusi akademik. Bagi pengembang aplikasi, analisis sentimen secara otomatis dapat menjadi alat penting untuk mengidentifikasi masalah yang sering dihadapi pengguna, seperti keluhan terhadap performa aplikasi, kesulitan dalam proses pemesanan tiket, atau ketidakpuasan terhadap tampilan antarmuka. Dengan memahami sentimen pengguna, pengembang dapat mengambil keputusan yang lebih tepat dan cepat dalam melakukan perbaikan aplikasi.

Bagi peneliti, studi ini menunjukkan bahwa karakteristik data sangat mempengaruhi performa algoritma. Hal ini menegaskan pentingnya proses exploratory data analysis sebelum memilih model. Penelitian serupa dapat dilakukan pada aplikasi digital lainnya untuk membandingkan hasil dan memperluas cakupan studi.

Dari sisi akademik, penelitian ini memberikan kontribusi terhadap literatur dalam bidang analisis sentimen berbahasa Indonesia, yang masih tergolong kurang jika dibandingkan dengan penelitian berbahasa Inggris. Pendekatan supervised learning yang digunakan juga dapat dijadikan acuan dalam pengembangan sistem rekomendasi, chatbot, dan sistem evaluasi berbasis opini.

3.9. Visualisasi Hasil Klasifikasi

Untuk memperkuat pemahaman terhadap performa model klasifikasi, ditampilkan confusion matrix untuk masing-masing model. Confusion matrix memberikan gambaran detail mengenai jumlah prediksi

yang benar dan salah yang dilakukan oleh model terhadap masing-masing kelas. Matriks ini merupakan alat visualisasi penting dalam evaluasi kinerja karena memungkinkan analisis lebih dalam terhadap kesalahan klasifikasi yang terjadi.

3.9.1. Confusion Matrix Naïve Bayes

Tabel 1. Confusion Matrix Hasil Klasifikasi Naïve Bayes

	Prediksi Positif	**Prediksi Negatif**
Aktual Positif	417	43
Aktual Negatif	45	415

Berdasarkan matriks di atas, model Naïve Bayes berhasil mengklasifikasikan 417 dari 460 data aktual positif secara benar, sementara 43 lainnya salah diklasifikasikan sebagai negatif. Untuk data aktual negatif, sebanyak 415 berhasil diklasifikasikan dengan benar, dan 45 sisanya salah prediksi sebagai positif. Dari total 920 data uji, model hanya melakukan 88 kesalahan klasifikasi (false positive dan false negative), menunjukkan performa yang sangat baik dengan distribusi kesalahan yang relatif seimbang.

3.9.2. Confusion Matrix Support Vector Machine

Tabel 2. Confusion Matrix Hasil Klasifikasi Support Vector Machine (SVM)

	Prediksi Positif	**Prediksi Negatif**
Aktual Positif	408	52
Aktual Negatif	51	409

Model SVM juga menunjukkan hasil yang solid, meskipun sedikit lebih rendah dibandingkan Naïve Bayes. Sebanyak 408 data aktual positif berhasil diklasifikasikan dengan benar, sedangkan 52 diklasifikasikan secara keliru. Untuk data aktual negatif, 409 data berhasil dikenali dengan tepat dan 51 lainnya mengalami kesalahan klasifikasi. Total kesalahan klasifikasi sebanyak 103, atau 11,2% dari data uji. Meskipun demikian, SVM tetap menghasilkan kinerja yang baik, hanya sedikit tertinggal dari Naïve Bayes.

Visualisasi ini penting untuk menyoroti bahwa baik Naïve Bayes maupun SVM memiliki kemampuan generalisasi yang tinggi terhadap data uji. Namun, model Naïve Bayes memiliki tingkat kesalahan yang sedikit lebih rendah, sekaligus menunjukkan kestabilan prediksi antar kelas.

3.10. Analisis Kesalahan (Error Analysis)

Analisis kesalahan atau error analysis bertujuan untuk mengidentifikasi jenis-jenis kesalahan yang terjadi selama proses klasifikasi dan memahami penyebabnya. Dalam konteks analisis sentimen berbasis teks pendek seperti ulasan aplikasi, terdapat beberapa faktor yang menyebabkan kesalahan prediksi.

Salah satu penyebab utama adalah keberadaan ulasan yang mengandung ambiguitas semantik atau sarkasme. Ulasan dengan sarkasme sering kali menggunakan kata-kata positif secara eksplisit, namun dalam konteks sebenarnya bermakna negatif. Contohnya adalah ulasan berikut: “Mantap, gagal terus waktu mau checkout. Aplikasi terbaik, ya buat emosi!”. Ulasan ini secara literal menggunakan kata-kata positif seperti “mantap” dan “terbaik”, yang menyebabkan model kesulitan dalam mengidentifikasi bahwa ulasan tersebut sebenarnya bernada negatif. Model klasifikasi yang hanya mengandalkan frekuensi kata akan mudah tertipu oleh sarkasme karena tidak memahami konteks emosional atau nada ironi.

Selain itu, kesalahan juga disebabkan oleh penggunaan kata-kata yang tidak baku atau kata serapan yang tidak dikenali oleh proses stemming atau normalisasi. Misalnya, kata “laggy”, “buffering” atau “crash” kadang luput dari kamus normalisasi sehingga tidak dipetakan ke dalam fitur yang relevan. Hal ini menunjukkan pentingnya pengembangan kamus normalisasi dan stopword yang lebih komprehensif, khususnya untuk Bahasa Indonesia yang banyak mengandung variasi slang dan singkatan informal.

Beberapa ulasan juga memiliki struktur kalimat yang kompleks atau menyertakan lebih dari satu opini, misalnya: “Aplikasinya bagus, tapi loading-nya lama dan pembayaran sering error”. Dalam kasus seperti ini, sentimen bercampur (mixed sentiment) membuat model sulit menentukan label secara tepat karena ulasan mengandung elemen positif dan negatif sekaligus.

3.11. Ringkasan Temuan

Berdasarkan seluruh proses dan hasil evaluasi yang telah dibahas, dapat dirumuskan beberapa temuan utama yang mencerminkan kontribusi signifikan dari penelitian ini. Pertama, algoritma Naïve Bayes

terbukti memiliki kinerja yang lebih unggul dibandingkan Support Vector Machine dalam klasifikasi sentimen terhadap ulasan pengguna aplikasi CGV. Hal ini terlihat dari hasil akurasi yang mencapai 90,38% dan nilai F1-score sebesar 90,46%, lebih tinggi dibandingkan SVM yang memiliki akurasi 88,88% dan F1-score 89,08%. Hasil ini menegaskan bahwa pendekatan probabilistik seperti Naïve Bayes sangat efektif untuk klasifikasi teks pendek yang bersifat eksplisit dan tidak kompleks.

Kedua, penggunaan teknik preprocessing dan representasi fitur TF-IDF dengan konfigurasi ngram yang sesuai memberikan kontribusi besar terhadap peningkatan akurasi model. Kombinasi antara unigram dan bigram, serta pemilihan kata dengan frekuensi optimal (min_df=5 dan max_df=0.95) terbukti mampu menyaring fitur yang benar-benar relevan.

Ketiga, hasil analisis juga menunjukkan bahwa salah satu tantangan terbesar dalam klasifikasi sentimen berbasis teks adalah keberadaan ulasan ambigu, campuran, atau bersifat sarkastik. Untuk mengatasi tantangan ini, penelitian selanjutnya dapat mengeksplorasi penggunaan pendekatan berbasis deep learning, seperti LSTM atau BERT, yang memiliki kemampuan memahami konteks dan urutan kata secara lebih mendalam.

Keempat, dari segi efisiensi, Naïve Bayes memiliki keunggulan dalam waktu pelatihan dan kemudahan implementasi, menjadikannya pilihan ideal dalam aplikasi nyata yang membutuhkan analisis cepat dan berskala besar.

Terakhir, penelitian ini memberikan bukti nyata bahwa analisis sentimen terhadap ulasan pengguna dapat menjadi alat evaluasi yang efektif bagi pengembang aplikasi. Informasi yang diperoleh dari model klasifikasi dapat digunakan untuk mendeteksi permasalahan utama yang dihadapi pengguna dan merancang strategi perbaikan yang lebih tepat sasaran.

4. KESIMPULAN

Penelitian ini bertujuan untuk menganalisis sentimen pengguna terhadap aplikasi CGV Cinemas Indonesia yang diulas melalui platform Google Play Store, serta membandingkan performa dua algoritma klasifikasi teks yang populer, yaitu Naïve Bayes dan Support Vector Machine (SVM). Melalui serangkaian tahapan mulai dari pengumpulan data, pra-pemrosesan teks, penyeimbangan data, hingga ekstraksi fitur menggunakan metode TF-IDF, penelitian ini berhasil mengelola dan mengolah 4.600 data ulasan yang terbagi rata antara sentimen positif dan negatif. Hasil eksperimen menunjukkan bahwa Naïve Bayes memiliki performa yang lebih unggul dibandingkan SVM dalam hal akurasi dan efisiensi komputasi. Naïve Bayes mencatat akurasi sebesar 90,38% dan nilai F1-score 90,46%, sedangkan SVM berada sedikit di bawahnya dengan akurasi 88,88% dan F1-score 89,08%. Keunggulan ini menandakan bahwa model probabilistik seperti Naïve Bayes sangat sesuai untuk menangani teks pendek dengan struktur sederhana seperti ulasan aplikasi. Selain lebih akurat, model ini juga lebih ringan dan cepat dalam proses pelatihan, sehingga ideal digunakan dalam sistem klasifikasi otomatis berbasis sumber daya terbatas. Dengan demikian, hasil dari penelitian ini diharapkan dapat menjadi rujukan bagi pengembang aplikasi dalam mengevaluasi kualitas layanan berdasarkan opini pengguna. Untuk pengembangan lebih lanjut, disarankan agar penelitian berikutnya mengintegrasikan teknik Natural Language Processing (NLP) lanjutan seperti word embedding (misalnya Word2Vec, GloVe, atau BERT), serta mempertimbangkan analisis terhadap sentimen netral dan ulasan bersifat sarkastik guna meningkatkan ketepatan dan ketahanan model dalam menghadapi berbagai variasi bahasa alami.

REFERENSI

[1] Apriliyani, M., Musyaffaq, M. I., Nur'Aini, S., Handayani, M. R., & Umam, K. (2024). Implementasi analisis sentimen pada ulasan aplikasi Duolingo di Google Playstore menggunakan algoritma Naïve Bayes. *AITI*, *21*(2), 298–311. https://doi.org/10.24246/aiti.v21i2.298-311

[2] Astuti, K. C., Firmansyah, A., & Riyadi, A. (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://doi.org/http://doi.org/10.33395/remik.v8i1.13421

[3] Asyhari, M. Y., Juwari, J., Hapsari, E. D., & Yulianto, S. (2023). Pendekatan Metode Kolokasi untuk Text Processing Ulasan Aplikasi Android Surveilans Penyebaran Covid-19 di Indonesia. *Journal Information System Development (ISD)*, *8*(1), 33–42. https://doi.org/10.19166/isd.v8i1.586

[4] Betesda, B., Purwanto, H., Nuryadi, H., Sinaga, D., Al Din, S. J., & Al Afghani, D. Y. (2024). ANALISA SENTIMEN DATA ULASAN PADA GOOGLE PLAY DENGAN MENGGUNAKAN ALGORITMA NAÏVE BAYES DAN SUPPORT VECTOR MACHINE. *Jurnal Sains Dan Teknologi ISTP*, *22*(01), 08–15. https://doi.org/10.59637/jsti.v22i01.423

[5] Eldo, H., Ayuliana, A., Suryadi, D., Chrisnawati, G., & Judijanto, L. (2024). Penggunaan Algoritma

Support Vector Machine (SVM) Untuk Deteksi Penipuan pada Transaksi Online. *Jurnal Minfo Polgan*, *13*(2), 1627–1632. https://doi.org/10.33395/jmp.v13i2.14186

[6] Fazrian, V., Suprapti, T., & Narasati, R. (2024). PENERAPAN ALGORITMA NAIVE BAYES TERHADAP ANALISIS SENTIMEN APLIKASI GAME MULTIPLAYER ONLINE BATTLE ARENA. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *8*(1), 1005–1012. https://doi.org/10.36040/jati.v8i1.8432

[7] Gumilar, R. B., Cahyana, Y., Sukmawati, C. E., & Siregar, A. M. (2024). Analisa Perbandingan Algoritma Support Vector Machine dan K-Nearest Neighbors Terhadap Ulasan Aplikasi Vidio. *Journal of Information System Research (JOSH)*, *5*(4), 1188–1195. https://doi.org/10.47065/josh.v5i4.5640

[8] Haq, M. Z., Octiva, C. S., Ayuliana, A., Nuryanto, U. W., & Suryadi, D. (2024). Algoritma Naïve Bayes untuk Mengidentifikasi Hoaks di Media Sosial. *Jurnal Minfo Polgan*, *13*(1), 1079–1084. https://doi.org/10.33395/jmp.v13i1.13937

[9] Java, M. A., Mohammad Syafrullah, Windarto, W., & Painem, P. (2024). Analisis Sentimen Ulasan Pengguna Aplikasi Threads pada Google Play Store Menggunakan Multinomial Naive Bayes dan Support Vector Machine. *Jurnal Ticom: Technology of Information and Communication*, *12*(2), 75–80. https://doi.org/10.70309/ticom.v12i2.112

[10] Lubis, A. Y., & Setyawan, M. Y. H. (2024). Analisis Sentimen Terhadap Aplikasi Pospay Menggunakan Algoritma Support Vector Machine dan Naive Bayes. *Jurnal Teknologi Dan Sistem Informasi Bisnis*, *6*(3), 514–521. https://doi.org/10.47233/jteksis.v6i3.1310

[11] Maarif, M. M., & Setiyawati, N. (2024). Analisis Sentimen Review Aplikasi LinkedIn di Google Play Store Menggunakan Support Vector Machine. *Progresif: Jurnal Ilmiah Komputer*, *20*(1), 454. https://doi.org/10.35889/progresif.v20i1.1614

[12] 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

[13] Sari, P. K., & Suryono, R. R. (2024). KOMPARASI ALGORITMA SUPPORT VECTOR MACHINE DAN RANDOM FOREST UNTUK ANALISIS SENTIMEN METAVERSE. *Jurnal Mnemonic*, *7*(1), 31–39. https://doi.org/10.36040/mnemonic.v7i1.8977

[14] Sari, R. Y., Subandi, A., & Irsyad, I. (2024). Pengaruh Penggunaan Sistem Informasi Manajemen Berbasis Digital Terhadap Efisiensi Administrasi Pendidikan. *Academy of Social Science and Global Citizenship Journal*, *4*(1), 21–29. https://doi.org/10.47200/aossagcj.v4i1.2389

[15] Xu, W., Chen, J., Ding, Z., & Wang, J. (2024). Text sentiment analysis and classification based on bidirectional Gated Recurrent Units (GRUs) model. *Applied and Computational Engineering*, *77*(1), 132–137. https://doi.org/10.54254/2755-2721/77/20240670

[16] Yusuf Rismanda Gaja, M., Maulana, I., & Komarudin, O. (2024). ANALISIS SENTIMEN OPINI PENGGUNA APLIKASI VIDIO PADA ULASAN PLAYSTORE MENGGUNAKAN ALGORITMA NAIVE BAYES. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *7*(4), 2767–2774. https://doi.org/10.36040/jati.v7i4.7197