

ANALISIS KOMPARATIF ALGORITMA NAIVE BAYES DAN SVM PADA KLASIFIKASI SENTIMEN ULASAN NETFLIX

Alisa Hasna Khaira Aswaha, Moh. Muhtarom, Pramono

Fakultas Ilmu Komputer, Universitas Duta Bangsa Surakarta

Jl. Bhayangkara No.55, Tipes, Kec. Serengan, Kota Surakarta, Jawa Tengah 57154, Indonesia

220103204@mhs.udb.ac.id

ABSTRAK

Netflix sebagai layanan streaming digital menghasilkan jutaan ulasan pengguna di Google Play Store yang memuat opini terhadap kualitas layanan. Volume ulasan yang sangat besar membuat analisis manual tidak efektif dan rentan subjektivitas, sehingga dibutuhkan metode klasifikasi otomatis untuk mengelompokkan sentimen pengguna ke dalam kategori positif, netral, dan negatif. Penelitian ini bertujuan menganalisis dan membandingkan kinerja algoritma Naive Bayes dan Support Vector Machine (SVM) dalam klasifikasi sentimen ulasan pengguna aplikasi Netflix. Penelitian menggunakan pendekatan text mining dengan data sekunder berupa 110.418 ulasan dari platform Kaggle. Pelabelan sentimen dilakukan berdasarkan rating pengguna, di mana score 1–2 dikategorikan negatif, score 3 netral, dan score 4–5 positif. Tahapan penelitian meliputi preprocessing data, transformasi TF-IDF, klasifikasi, serta evaluasi menggunakan accuracy, precision, recall, dan F1-score. Hasil penelitian menunjukkan SVM memiliki kinerja lebih baik dibandingkan Naive Bayes, dengan accuracy 78,48% dan F1-score 75,34%, berbanding accuracy 77,30% dan F1-score 73,23% pada Naive Bayes. Kedua algoritma mengalami kesulitan mengklasifikasikan sentimen netral akibat ketidakseimbangan data, di mana kelas netral hanya mencakup 10,66% dari total dataset. SVM direkomendasikan sebagai algoritma yang lebih optimal untuk kasus klasifikasi sentimen serupa.

Kata kunci : Analisis sentimen, Naive Bayes, SVM, Netflix, Text Mining, TF-IDF

1. PENDAHULUAN

Netflix tercatat sebagai salah satu layanan streaming mencapai lebih dari 230 juta pelanggan aktif yang tersebar di berbagai negara, termasuk Indonesia[1].

Besarnya jumlah pengguna ini secara langsung berdampak pada melimpahnya data ulasan yang diunggah pengguna melalui platform seperti Google Play Store, yang memuat beragam opini, pengalaman, dan kritik terhadap layanan Netflix, mulai dari kualitas konten, kecepatan aplikasi, hingga pengalaman berlangganan.

Ulasan pengguna merupakan sumber data strategis yang dapat dimanfaatkan oleh pengembang untuk memahami kepuasan pengguna, mengidentifikasi kelemahan layanan, dan merancang strategi peningkatan kualitas produk. Akan tetapi, besarnya volume data ulasan yang terus meningkat secara eksponensial menjadi sebuah hambatan tersendiri yang harus dihadapi. Analisis secara manual tidak lagi efektif dan efisien untuk dilakukan dalam skala besar, mengingat proses tersebut membutuhkan waktu, tenaga, dan sumber daya yang tidak sedikit, serta rentan terhadap subjektivitas penilai. Oleh karena itu, dibutuhkan pendekatan otomatis yang mampu memproses dan mengklasifikasikan data ulasan secara cepat dan akurat.

Guna mengelompokkan opini publik secara otomatis, teknik Natural Language Processing (NLP) menyediakan metode analisis sentimen. Pendekatan ini terbukti ampuh dalam memetakan respons pengguna ke dalam kategori positif, netral, maupun negatif. Dengan cara ini, pengembang bisa langsung

memantau bagaimana respons pengguna saat itu juga. Alhasil, kebijakan atau langkah selanjutnya dapat diambil secara lebih cepat karena didukung oleh data yang akurat. Dari sekian banyak algoritma untuk analisis sentimen, Naive Bayes dan Support Vector Machine (SVM) menjadi dua metode yang paling sering diteliti dan diuji cobakan. Naive Bayes dikenal karena proses komputasinya yang ringan dan mudah diimplementasikan, sementara SVM diandalkan untuk mengolah data teks dengan dimensi yang kompleks, sehingga akurasi atau hasil klasifikasinya cenderung jauh lebih memuaskan[2].

Sayangnya, studi komparatif yang mengadu kedua metode tersebut pada sektor ulasan platform streaming seperti Netflix sejauh ini jumlahnya masih sangat minim. Kesenjangan inilah yang menjadi landasan rasionalisasi dilakukannya penelitian ini, yakni untuk mengisi kekosongan kajian ilmiah sekaligus memberikan rekomendasi algoritmik yang tepat bagi pengembangan sistem analisis sentimen pada domain sejenis.

Berdasarkan permasalahan dan rasionalisasi yang telah diuraikan, Fokus penelitian ini adalah mengevaluasi akurasi Naive Bayes dan SVM dalam mengklasifikasikan sentimen pengguna Netflix di Google Play Store. Dengan melihat nilai accuracy hingga F1-score, studi ini diharapkan dapat menunjukkan algoritma mana yang paling pas untuk memproses data ulasan di platform streaming digital[3].

Tahap pertama adalah pengumpulan data, yaitu mengambil dataset ulasan pengguna Netflix dari platform Kaggle yang bersumber dari Google Play

Store. Langkah berikutnya difokuskan pada tahap preprocessing data yang meliputi pembersihan karakter (cleaning), penyamaan format huruf (case folding), pemotongan teks (tokenization), penghapusan kata tidak penting (stopword removal), hingga pencarian kata dasar (stemming). Selanjutnya, representasi teks diubah ke bentuk angka numerik berbobot menggunakan teknik TF-IDF agar dapat diolah di tahap machine learning.

Memasuki proses klasifikasi sentimen, Naive Bayes bekerja mengandalkan perhitungan probabilitas Bayes, sementara SVM mencari hyperplane paling optimal untuk memisahkan kategori sentimen dengan margin tertinggi. Rangkaian riset ini ditutup dengan tahap evaluasi, di mana performa klasifikasi ulasan aplikasi Netflix dari kedua model diukur menggunakan indikator accuracy, precision, recall, serta F1-score.

2. TINJAUAN PUSTAKA

Kajian mengenai analisis sentimen pada ulasan aplikasi digital terus berkembang seiring pentingnya mengetahui respons pengguna terhadap platform berbasis teknologi. Dalam konteks aplikasi Netflix, Ramadani dkk. mengomparasi efektivitas algoritma Support Vector Machine (SVM), Decision Tree, dan Logistic Regression. Dari pengujian tersebut, SVM terbukti mengungguli metode lainnya dengan capaian akurasi 88,18%, tingkat precision 87,5%, recall 86,2%, serta F1-score 86,8%. Melalui metode K-Fold Cross Validation, SVM juga memperlihatkan tingkat stabilitas yang tinggi, menegaskan keandalannya dalam memproses data tekstual berskala besar[4].

Pengujian performa SVM juga diteliti oleh Maulana dkk[5]. pada ranah aplikasi investasi digital Pluang, yang mempertemukannya dengan algoritma Naive Bayes. Hasilnya, SVM kembali menunjukkan dominasinya dengan meraih akurasi 99,50%, precision 99,67%, recall 99,33%, dan F1-score 99,50%, sedikit di atas Naive Bayes yang mencatatkan akurasi 99,25%. Keunggulan SVM di sini didorong oleh kemampuan teknik kernel dalam memetakan data ke ruang dimensi yang lebih tinggi, sehingga batas antara sentimen positif, netral, maupun negatif menjadi lebih tegas.

Di sisi lain, perbandingan kedua metode ini pada data media sosial Twitter terkait tren metaverse juga pernah diangkat oleh Parameswari dkk[6]. Studi tersebut mengadu fungsi LinearSVC dari SVM dengan MultinomialNB milik Naive Bayes. Berdasarkan hasil evaluasi, SVM unggul dengan akurasi 78,3% dan Macro-F1 72,8%, berbanding terbalik dengan Naive Bayes yang hanya meraih akurasi 72,4% serta Macro-F1 60,2%. Riset ini sekaligus mengonfirmasi bahwa SVM lebih seimbang dan jeli dalam mengidentifikasi seluruh spektrum sentimen, khususnya ulasan yang bernada negatif.

Kendati sering berada di bawah bayang-bayang SVM, Naive Bayes tetap menunjukkan performa yang kompetitif. Hal ini dibuktikan oleh Wahyuni[7]. saat

menganalisis opini masyarakat mengenai isu sosial feminisme di Twitter. Algoritma ini sukses menghasilkan tingkat akurasi sebesar 87%, precision 89%, recall 86%, dan F1-score 86%. Temuan ini memperkuat alasan bahwa Naive Bayes tetap menjadi opsi yang relevan untuk klasifikasi teks berkat proses komputasinya yang ringkas dan efisien.

Berdasarkan penelitian terdahulu tersebut, dapat disimpulkan bahwa algoritma Naive Bayes dan SVM memang menjadi dua pilar algoritma yang paling diandalkan dalam studi analisis sentimen teks. Sayangnya, mayoritas riset yang ada saat ini masih didominasi oleh objek data dari media sosial seperti Twitter atau platform digital di luar industri hiburan. Studi komparasi yang berfokus membedah ulasan aplikasi Netflix khususnya yang bersumber dari Google Play Store masih relatif jarang ditemukan. Gap penelitian inilah yang ingin diisi melalui studi ini, yaitu dengan menguji dan membandingkan performa kedua algoritma tersebut berdasarkan pada metrik evaluasi accuracy, precision, recall, hingga F1-score.

3. METODE PENELITIAN

3.1. Tahapan Penelitian



Gambar Diagram 1. Diagram Alir Tahapan Penelitian Analisis Sentimen Netflix

Tahapan penelitian dalam studi ini disusun secara sistematis untuk memberikan gambaran alur proses penelitian yang dilakukan. Penelitian diawali dengan identifikasi masalah, dilanjutkan dengan pengumpulan data ulasan pengguna Netflix, preprocessing data teks,

transformasi data menggunakan metode TF-IDF, proses klasifikasi menggunakan algoritma Naive Bayes dan Support Vector Machine (SVM), evaluasi model, serta analisis hasil perbandingan performa kedua algoritma.

3.2. Identifikasi Masalah

Tahap ini berfokus pada analisis dan perbandingan performa algoritma Naive Bayes dan Support Vector Machine (SVM) dalam mengklasifikasikan sentimen ulasan pengguna aplikasi Netflix. Model dievaluasi menggunakan metrik accuracy, precision, recall, dan F1-score untuk menentukan algoritma dengan performa terbaik.

3.3. Pengumpulan Data

Dataset yang digunakan diambil dari platform Kaggle. Data tersebut pada dasarnya berasal dari kumpulan ulasan yang ditinggalkan pengguna aplikasi Netflix di Google Play Store. Isi dari dataset ini mencakup dua hal pokok, yaitu komentar tertulis dari pengguna dan angka rating yang mereka berikan.

3.4. Pra-Pemrosesan Data

Data teks yang baru dikumpulkan umumnya masih dalam kondisi mentah dan belum siap untuk langsung dianalisis. Oleh karena itu, diperlukan serangkaian proses pembersihan dan persiapan data yang disebut pra-pemrosesan. Berikut adalah tahapan yang dilakukan:

- Cleaning, pada tahap ini dilakukan pembersihan menyeluruh terhadap data teks dari berbagai elemen yang tidak diperlukan, seperti tanda baca, angka, simbol, URL, maupun emoji.
- Case Folding, seluruh teks diubah menjadi huruf kecil secara seragam. Hal ini dilakukan agar kata yang sama namun ditulis dengan format berbeda tetap dikenali sebagai satu kata yang identik oleh sistem.
- Tokenization, teks yang sudah dibersihkan kemudian dipecah menjadi satuan-satuan kata yang lebih kecil atau disebut token.
- Stopward Removal, tidak semua kata dalam sebuah kalimat memiliki makna yang penting.
- Stemming, Mengembalikan setiap kata ke bentuk dasarnya

3.5. Transformasi Data Menggunakan TF-IDF

Data teks dikonversi menjadi bentuk numerik menggunakan metode TF-IDF (Term Frequency – Inverse Document Frequency). Metode ini memberikan bobot tinggi pada kata yang spesifik (jarang muncul di ulasan lain) dan bobot rendah pada kata umum yang kurang informatif. Hasil pembobotan angka ini kemudian digunakan sebagai input untuk klasifikasi pada algoritma Naive Bayes dan SVM.

Konsep dasar TF-IDF dapat dituliskan sebagai berikut:

$$TF - IDF(t, d) = TF(t, d) \times IDF(t) \quad (1)$$

3.6. Klasifikasi Menggunakan Naive Bayes

Dataset representasi TF-IDF dibagi menjadi data latih (untuk mengenali pola) dan data uji (untuk mengukur kemampuan klasifikasi). Proses ini mengklasifikasikan ulasan ke dalam tiga kelompok: positif, netral, dan negatif. Naive Bayes bekerja dengan menghitung peluang kemunculan setiap kata pada masing-masing kategori sentimen, di mana kategori dengan nilai peluang tertinggi dipilih sebagai hasil prediksi akhir berdasarkan Teorema Bayes:

$$P(C|X) = \frac{P(X|C)P(C)}{P(X)} \quad (2)$$

3.7. Klasifikasi Menggunakan Support Vector Machine (SVM)

Sebagai pembanding, digunakan algoritma Support Vector Machine (SVM) yang bekerja dengan menemukan batas pemisah paling optimal (hyperplane) antar kelas. Model dilatih menggunakan data latih hasil transformasi TF-IDF, lalu diuji untuk melihat kemampuannya mengelompokkan ulasan ke dalam kategori positif, netral, dan negatif. Hasil kedua model nantinya dibandingkan untuk menentukan algoritma terbaik.

3.8. Evaluasi Kinerja Model

Evaluasi kinerja model dilakukan menggunakan confusion matrix dengan empat metrik utama:

- Accuracy, Proporsi data yang berhasil diklasifikasikan dengan benar dari keseluruhan data uji.
- Precision, Ketepatan model ketika memberikan prediksi kelas positif.
- Recall, Kemampuan model dalam menemukan seluruh data yang seharusnya masuk kategori positif.
- F1-score, Nilai gabungan yang menyeimbangkan antara precision dan recall.

3.9. Hasil Evaluasi dan Perbandingan

Performa Naive Bayes dan SVM dianalisis berdasarkan keempat metrik di atas untuk menentukan metode klasifikasi paling optimal. Kategori sentimen ditentukan berdasarkan nilai rating pengguna di Google Play Store:

- Positif (Rating 4-5) : Mengandung kepuasan atau pujian terhadap kualitas konten dan fitur aplikasi (misal: "Great app, love the content").
- Negatif (Rating 1-2) : Mengandung keluhan teknis seperti bug atau aplikasi crash (misal: "The app keeps crashing"). Kategori ini mendominasi dataset dengan 53.822 data (48,74%).
- Positif (Rating 3) : Bersifat campuran antara apresiasi dan kritik tanpa emosi dominan (misal: "I like the app but there's a bug"). Kategori ini memiliki data paling sedikit, yaitu 11.772 data (10,66%), yang memicu ketidakseimbangan kelas (class imbalance) sebagai tantangan bagi kedua algoritma.

4. HASIL DAN PEMBAHASAN

4.1. Hasil Pengumpulan Data

Data utama dalam penelitian ini adalah 110.418 baris ulasan pengguna aplikasi Netflix dari Google Play Store (via Kaggle) dengan format CSV. Dataset ini memuat 8 atribut: reviewId, userName, content, score, thumbsUpCount, reviewCreatedVersion, at, dan appVersion. Atribut content digunakan sebagai sumber data teks, sedangkan score digunakan sebagai acuan label sentimen. Data ini bersifat tidak terstruktur sehingga membutuhkan tahap preprocessing sebelum diklasifikasikan dengan algoritma Naive Bayes dan Support Vector Machine. Contoh data ulasan disajikan pada Tabel 1.

Tabel 1. Data Ulasan Pengguna Netflix

No	userName	Content	Score
1	Evans William	Good	5
2	Issa Tunda	Thank you	5
3	Gisa Brian	So smart and has translated movies	4
4	Devon Bryant	Won't open on my phone....	1

No	userName	Content	Score
5	Yshmael Galinato	I like the app but there's a bug tha is changing the language of my Netflix to Chinese and it's very irritating	3

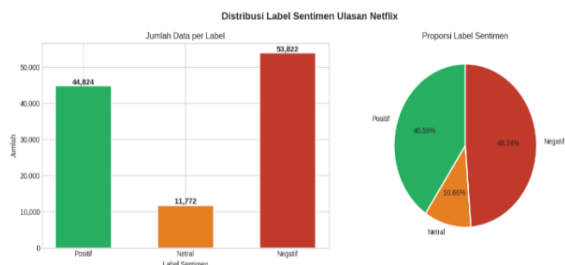
4.2. Pelabelan Sentimen

Dalam penelitian ini, pelabelan dilakukan secara otomatis berdasarkan nilai score atau rating yang diberikan pengguna saat menulis ulasan di Google Play Store. Sentimen dibagi ke dalam tiga kelompok utama sebagai berikut:

Ulasan dengan rating 4 dan 5 dikategorikan sebagai sentimen positif, karena nilai tersebut mencerminkan kepuasan pengguna terhadap aplikasi Netflix. Ulasan dengan rating 3 masuk ke dalam sentimen netral, karena cenderung bersifat campuran antara apresiasi dan kritik tanpa kecenderungan yang dominan ke salah satu arah. Sementara itu, ulasan dengan rating 1 dan 2 dikategorikan sebagai sentimen negatif, karena menggambarkan ketidakpuasan atau keluhan pengguna terhadap layanan yang diterima.

Tabel 2. Pelabelan Sentimen

No	userName	Content	Score	Label Sentimen
1	Evans William	Good	5	Positif
2	Issa Tunda	Thank you	5	Positif
3	Gisa Brian	So smart and has translated movies	4	Positif
4	Devon Bryant	Won't open on my phone....	1	Negatif
5	Yshmael Galinato	I like the app but there's a bug tha is changing the language of my Netflix to Chinese and it's very irritating	3	Netral



Gambar Diagram 2. Distribusi Label Sentimen Ulasan Netflix

Berdasarkan hasil pelabelan pada Gambar 2, distribusi data terbagi menjadi: sentimen positif bermasalah 44.824 data (40,59%), sentimen netral 11.772 data (10,66%), dan sentimen negatif

mendominasi dengan 53.822 data (48,74%). Distribusi tersebut memperlihatkan bahwa keluhan dan kritik pengguna mendominasi dataset, yang mengindikasikan masih banyaknya permasalahan teknis, bug, maupun error pada aplikasi.

4.3. Cleaning dan Case Folding

Proses case folding dilakukan terlebih dahulu dengan mengubah seluruh karakter dalam teks menjadi huruf kecil agar sistem tidak memperlakukan kata yang sama sebagai dua kata berbeda. Selanjutnya, proses cleaning dilakukan menggunakan metode regular expression (regex) untuk membuang elemen yang tidak relevan seperti URL, tag HTML, emoji, angka, tanda baca, karakter khusus, dan spasi berlebih.

Tabel 3. Cleaning dan Case folding

Data Asli	Case Folding	Cleaning
Whenever i want to use netflix it keeps saying im not part of my family's netflix acc IM DELETING NETFLIX	whenever i want to use netflix it keeps saying im not part of my family's netflix acc im deleting netflix	whenever i want to use netflix it keeps saying im not part of my familys netflix acc im deleting netflix
Can't open my app anymore. Keeps flashing message playing on too many devices. I changed account. Still same.	can't open my app anymore. keeps flashing message playing on too many devices. i changed account. still same	cant open my app anymore keeps flashing message playing on too many devices i changed account still same

Data Asli	Case Folding	Cleaning
please bring back some old shows that got a lot of popularity. also Netflix tends to bring shows that had only 1 amazing season and then the show will never get another season.	please bring back some old shows that got a lot of popularity. also netflix tends to bring shows that had only 1 amazing season and then the show will never get another season.	please bring back some old shows that got a lot of popularity also netflix tends to bring shows that had only amazing season and then the show will never get another season

4.4. Tokenization

Proses ini bertujuan untuk memecah setiap kalimat dalam ulasan menjadi satuan kata yang lebih kecil (token) menggunakan fungsi `word_tokenize()`

dari library NLTK (Natural Language Toolkit). Tahap ini menjadi fondasi penting karena algoritma machine learning membutuhkan data dalam bentuk unit kata yang terpisah.

Tabel 4. Tokenization

Data Setelah Cleanning	Tokenization
whenever i want to use netflix it keeps saying im not part of my familys netflix acc im deleting netflix	[whenever, i, want, to, use, netflix, it, keeps, saying, im, not, part, of, my, familys, netflix, acc, im, deleting, netflix]
cant open my app anymore keeps flashing message playing on too many devices i changed account still same	[cant, open, my, app, anymore, keeps, flashing, message, playing, on, too, many, devices, i, changed, account, still, same]
please bring back some old shows that got a lot of popularity also netflix tends to bring shows that had only amazing season and then the show will never get another season	[please, bring, back, some, old, shows, that, got, a, lot, of, popularity, also, netflix, tends, to, bring, shows, that, had, only, amazing, season, and, then, the, show, will, never, get, another, season]

4.5. Stopward Removal dan Stemming

Stopword removal dilakukan dengan menghapus kata-kata umum dalam bahasa Inggris yang tidak membawa informasi sentimen signifikan (seperti "the", "to", "of"). Selanjutnya, proses stemming

dilakukan menggunakan algoritma Porter Stemmer untuk mengubah setiap kata ke bentuk dasarnya agar dimensi fitur menjadi lebih kecil dan proses klasifikasi menjadi lebih efisien.

Tabel 5. Stopward Removal dan Stemming

Data Setelah Tokenization	Stopword Removal	Stemming
[whenever, i, want, to, use, netflix, it, keeps, saying, im, not, part, of, my, familys, netflix, acc, im, deleting, netflix]	[whenever, want, use, netflix, keeps, saying, im, part, familys, netflix, acc, im, deleting, netflix]	whenev want use netflix keep say im part famili netflix acc im delet netflix
[cant, open, my, app, anymore, keeps, flashing, message, playing, on, too, many, devices, i, changed, account, still, same]	[cant, open, app, anymore, keeps, flashing, message, playing, many, devices, changed, account, still]	cant open app anymor keep flash messag play mani devic chang account still
[please, bring, back, some, old, shows, that, got, a, lot, of, popularity, also, netflix, tends, to, bring, shows, that, had, only, amazing, season, and, then, the, show, will, never, get, another, season]	[please, bring, back, old, shows, got, lot, popularity, also, netflix, tends, bring, shows, amazing, season, show, never, get, another, season]	pleas bring back old show got lot popular also netflix tend bring show amaz season show never get anoth season

Berdasarkan hasil yang ditampilkan pada Tabel 5, beberapa contoh perubahan yang terjadi antara lain kata "shows" menjadi "show", "popularity" menjadi "popular", dan "amazing" menjadi "amaz".

Secara keseluruhan, hasil akhir dari seluruh tahap preprocessing ini adalah data teks yang lebih ringkas, seragam, dan optimal untuk digunakan pada proses ekstraksi fitur menggunakan TF-IDF serta klasifikasi sentimen menggunakan Naive Bayes dan SVM.

4.6. Wordcloud

Visualisasi wordcloud digunakan untuk menampilkan kata-kata yang paling dominan dalam dataset, di mana ukuran kata sebanding dengan frekuensi kemunculannya. Berikut merupakan hasil visualisasi wordcloud dari data ulasan pengguna Netflix :



Gambar 3. Wordcloud Positif



Gambar 4. Wordcloud Negatif



Gambar 5. Wordcloud Netral

Berdasarkan hasil visualisasi pada Gambar 3, 4, dan 5, terlihat sejumlah kata yang menonjol dengan ukuran lebih besar dibandingkan yang lain. Kata-kata tersebut mencerminkan topik yang paling banyak dibicarakan oleh pengguna, mulai dari kualitas konten, performa aplikasi, hingga pengalaman penggunaan secara umum. Visualisasi ini memberikan pemahaman awal yang berguna sebelum masuk ke tahap analisis yang lebih mendalam.

4.7. TF - IDF

Tahap ini mengubah data teks menjadi representasi numerik melalui metode Term Frequency-Inverse Document Frequency (TF-IDF) karena algoritma machine learning hanya dapat menerima masukan berupa angka.

Tabel 6. Data Hasil TF - IDF

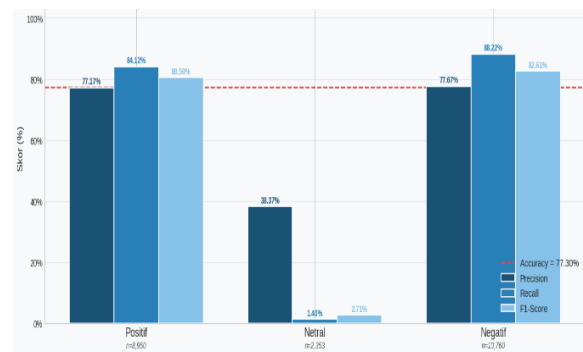
Kata	Bobot TF - IDF
app	2.0202
netflix	2.0746
watch	2.3615
show	2.5746
like	3.0454

TF-IDF memberikan nilai bobot pada setiap kata berdasarkan seberapa sering kata tersebut muncul dalam satu ulasan dibandingkan dengan kemunculannya di seluruh dataset. Kata yang spesifik akan mendapat bobot lebih tinggi, sedangkan kata

yang muncul hampir di semua ulasan akan mendapat bobot rendah.

4.8. Naive Bayes

Model Naive Bayes dilatih menggunakan data latih dan diuji pada data uji dengan menggunakan fitur masukan dari TF-IDF. Evaluasi dilakukan menggunakan classification report untuk mengukur kinerja model. Berikut merupakan hasil klasifikasi sentimen menggunakan algoritma Naive Bayes

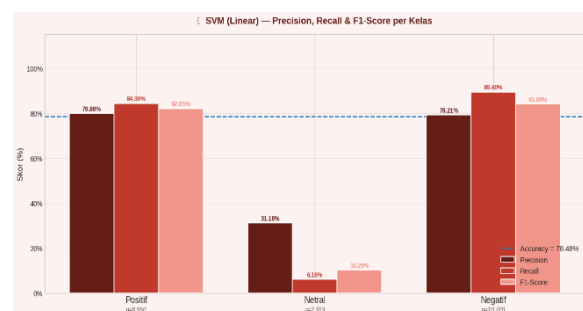


Gambar 6. Klasifikasi Naive Bayes

Model Naive Bayes dilatih menggunakan data latih dan diuji pada data uji dengan menggunakan fitur masukan dari TF-IDF. Evaluasi dilakukan menggunakan classification report untuk mengukur kinerja model.

4.9. Support Vector Machine (SVM)

SVM digunakan sebagai metode pembandingan yang bekerja dengan menemukan pemisah terbaik (hyperplane) antar kelas sentimen. SVM sangat baik dalam menangani data teks berdimensi besar hasil transformasi TF-IDF. Berikut merupakan hasil klasifikasi sentimen menggunakan algoritma Support Vector Machine (SVM) :

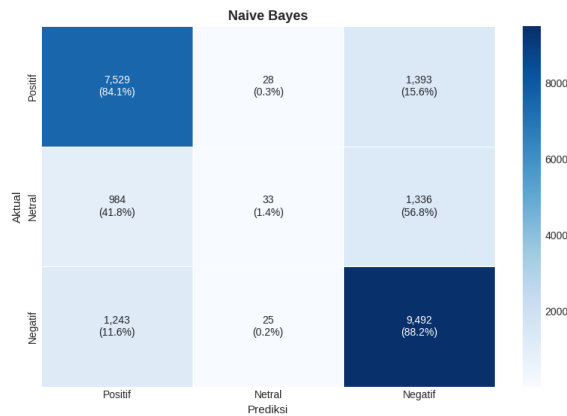


Gambar 7. Klasifikasi SVM

Mengacu pada Gambar 7, model SVM mencatatkan akurasi 78,48%, sedikit lebih tinggi dibandingkan Naive Bayes. SVM menunjukkan performa stabil pada kelas positif (F1-score 82,05%) dan performa terbaik pada kelas negatif (F1-score 83,99%). Kinerja pada kelas netral juga meningkat (F1-score 10,29%) dibandingkan Naive Bayes, meskipun masih belum optimal karena keterbatasan data netral.

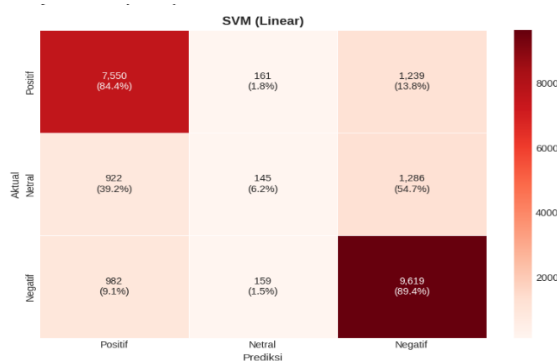
4.10. Confusion Matriks

Confusion matrix memberikan gambaran rinci mengenai jumlah prediksi tepat dan salah pada masing-masing kelas sentimen. Hasil confusion matrix ditampilkan pada Gambar 8 dan 9 berikut :



Gambar 8. Confusion Matriks Naïve Bayes

Pada model Naive Bayes, hasil confusion matrix adalah berhasil mengenali 84,1% data positif dan 88,2% data negatif dengan benar. Namun, pada kelas netral hanya 1,4% data yang terprediksi tepat, sedangkan sisanya dominan salah diprediksi sebagai negatif (56,8%) atau positif (41,8%).



Gambar 9. Confusion Matriks SVM

Menunjukkan peningkatan di mana 84,4% data positif dan 89,4% data negatif berhasil diklasifikasikan secara benar. Pada kelas netral, prediksi tepat meningkat menjadi 6,2%, membuktikan SVM lebih unggul dan stabil daripada Naive Bayes.

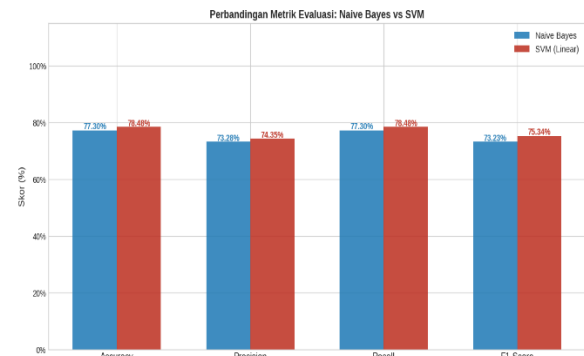
4.11. Evaluasi Model

Setelah dilakukan evaluasi secara pada masing-masing algoritma di subbab 4.8, 4.9 dan 4.10, maka pada subbab ini dilakukan perbandingan kinerja antara algoritma Naive Bayes dan Support Vector Machine (SVM) berdasarkan metrik evaluasi yang telah diperoleh. Perbandingan ini bertujuan untuk menentukan algoritma mana yang lebih efektif dalam mengklasifikasikan sentimen ulasan Netflix.

4.12. Perbandingan Metrik Evaluasi Keseluruhan

Tabel 7. Matriks Evaluasi Keseluruhan

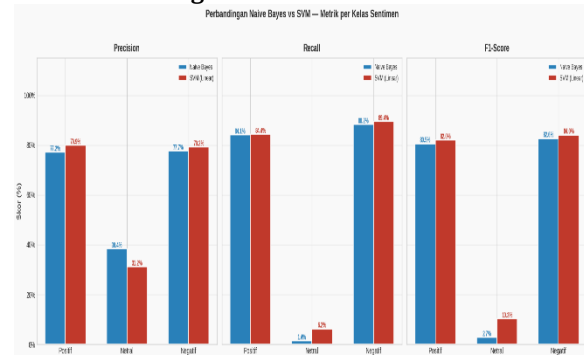
Metriks	Naive Bayes	SVM	Selisih
Accuracy	77.30%	78.48%	+1.18% (SVM)
Precision	73.28%	74.35%	+1.07% (SVM)
Recall	77.30%	78.48%	+1.18% (SVM)
F1-Score	73.23%	75.34%	+2.11% (SVM)



Gambar 10. Perbandingan Matriks Evaluasi Keseluruhan

Berdasarkan tabel di atas, Algoritma SVM secara konsisten mengungguli Naive Bayes pada seluruh metrik evaluasi keseluruhan dengan selisih performa tertinggi pada nilai F1-Score sebesar 2,11%.

4.13. Perbandingan Metrik Evaluasi Keseluruhan



Gambar 11. Perbandingan Matriks Evaluasi Keseluruhan

- Positif : SVM unggul di seluruh metrik (Precision 79,90%, Recall 84,40%, F1-Score 82,00%) dibandingkan Naive Bayes, menunjukkan SVM lebih andal dalam mengenali ulasan positif.
- Netral : Menjadi kelemahan utama kedua algoritma karena adanya *class imbalance* (ulasan netral jauh lebih sedikit). Meskipun Naive Bayes memiliki Precision lebih tinggi (38,40%), F1-Score SVM (10,30%) tetap lebih baik daripada Naive Bayes (2,70%) karena nilai *recall* yang lebih tinggi.
- Negatif : Merupakan kelas dengan performa terbaik bagi kedua model karena pola kata keluhan yang khas (seperti *crash*, *terrible*, *worst*). SVM kembali unggul dengan F1-Score 84,00% berbanding 82,60% pada Naive Bayes.

5. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa klasifikasi sentimen ulasan pengguna aplikasi Netflix dapat dilakukan secara efektif menggunakan pendekatan text mining dengan algoritma machine learning. Pelabelan sentimen dilakukan berdasarkan nilai score pengguna pada Google Play Store, di mana score 1–2 dikategorikan sebagai negatif, score 3 sebagai netral, dan score 4–5 sebagai positif, sehingga model dilatih untuk mengenali pola teks yang berkorelasi dengan masing-masing kategori tersebut.

Hasil evaluasi menunjukkan bahwa algoritma Support Vector Machine (SVM) memiliki kinerja lebih baik dibandingkan Naive Bayes, yang ditunjukkan oleh nilai accuracy 78,48% dan F1-score 75,34% berbanding accuracy 77,30% dan F1-score 73,23% pada Naive Bayes. SVM terbukti lebih mampu menangani karakteristik data ulasan yang tidak terstruktur dan memiliki variasi bahasa yang tinggi. Sementara itu, Naive Bayes meskipun lebih sederhana dan efisien secara komputasi, menunjukkan performa yang sedikit lebih rendah terutama dalam mengklasifikasikan sentimen netral.

Kedua algoritma mengalami kesulitan dalam mengenali kelas netral akibat ketidakseimbangan data, di mana sentimen netral hanya mencakup 10,66% dari total dataset. Hal ini disebabkan oleh sedikitnya pengguna yang memberikan score 3, sehingga model tidak memiliki cukup contoh untuk mempelajari pola sentimen netral secara optimal. Selain itu, pendekatan pelabelan berbasis rating memiliki keterbatasan karena isi teks ulasan tidak selalu mencerminkan nilai rating yang diberikan pengguna, sehingga penelitian selanjutnya disarankan menggunakan pelabelan manual atau pendekatan lexicon-based untuk menghasilkan label yang lebih akurat.

DAFTAR PUSTAKA

- [1] L. Rhomaningtias, A. Khairunisa, S. Shella, M. Wara, and K. M. Hindrayani, "ANALISIS SENTIMEN ULASAN APLIKASI SMILE INDONESIA MENGGUNAKAN METODE NAIVE BAYES DAN SUPPORT VECTOR MACHINE (SVM)", doi: 10.52972/hoaq.vol16no1.
- [2] D. Prasetia, N. Rahaningsih, R. D. Dana, and C. L. Rohmat, "ANALISIS SENTIMEN PENGGUNA APLIKASI MYBLUEBIRD DENGAN ALGORITMA NAÏVE BAYES DI PLAYSTORE," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 13, no. 1, Jan. 2025, doi: 10.23960/jitet.v13i1.5687.
- [3] D. P. Sari, R. Risman, F. Maulana, L. Efrizoni, and R. Rahmaddeni, "Model Prediksi Dampak Perubahan Iklim pada Ketahanan Pangan Menggunakan Algoritma Support Vector Machine and K-Nearest Neighbors," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 5, no. 3, pp. 851–861, Jun. 2025, doi: 10.57152/malcom.v5i3.1975.
- [4] N. C. Ramadani, I. Tahyudin, and A. Shouni Barkah, "Perbandingan Algoritma Support Vector Machine, Decision Tree, dan Logistic Regresion Pada Analisis Sentimen Ulasan Aplikasi Netflix," *Jurnal Nasional Teknologi dan Sistem Informasi*, vol. 10, no. 2, pp. 110–117, Aug. 2024, doi: 10.25077/teknosi.v10i2.2024.110-117.
- [5] B. A. Maulana, M. J. Fahmi, A. M. Imran, and N. Hidayati, "Analisis Sentimen Terhadap Aplikasi Pluang Menggunakan Algoritma Naive Bayes dan Support Vector Machine (SVM)," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 4, no. 2, pp. 375–384, Feb. 2024, doi: 10.57152/malcom.v4i2.1206.
- [6] S. D. Parameswari et al., "Studi Perbandingan Naïve Bayes dan Support Vector Machine (SVM) dalam Analisis Sentimen Pengguna Metaverse," *Jurnal Teknologi dan Manajemen Industri Terapan (JTMIT)*, vol. 4, no. 3, 2025.
- [7] W. Wahyuni, "Analisis Sentimen terhadap Opini Feminisme Menggunakan Metode Naive Bayes," *Jurnal Informatika Ekonomi Bisnis*, pp. 148–153, Sep. 2022, doi: 10.37034/infec.v4i4.162.
- [8] F. Fiddin and T. Hidayat, "Jurnal SAINTIKOM (Jurnal Sains Manajemen Informatika dan Komputer) Analisis Sentimen Ulasan Produk Sayur di Tokopedia Menggunakan Model Support Vector Machine Dengan Representasi TF-IDF", [Online]. Available: <https://ojs.trigunadharma.ac.id/index.php/jis/index>
- [9] S. F. Kadir and A. Fairuzabadi, "Analisis Sentimen Ulasan Shopee di Google Play dengan TF-IDF dan Logistic Regression," *RIGGS: Journal of Artificial Intelligence and Digital Business*, vol. 4, no. 2, pp. 7940–7945, Jul. 2025, doi: 10.31004/riggs.v4i2.2850.
- [10] A. Abidin, "ANALISIS SENTIMEN ULASAN APLIKASI SIREKAP MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 13, no. 3S1, Oct. 2025, doi: 10.23960/jitet.v13i3S1.7768