

LAPORAN

ANALISIS SENTIMEN KOMPARATIF PENGGUNA LAYANAN STREAMING NETFLIX DAN VIDIO

Dosen Pengampu: Akmal Setiawan Wijaya, S.Kom., M.Kom.



Disusun Oleh:

Vysca Rina Listianing Fitrianna

5220411025

**PROGRAM STUDI INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS TEKNOLOGI YOGYAKARTA**

2026

DAFTAR ISI

DAFTAR ISI.....	ii
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang	1
1.2. Tujuan	1
BAB II HASIL PEMBAHASAN.....	3
2.1. Akuisisi Data.....	3
2.2. Pra-pemrosesan Data.....	3
2.3. Metodologi Analisis	4
2.3.1. Analisis Deskriptif	4
2.3.2. Analisis Sentimen	4
2.4. Hasil	4
2.5. Pemodelan Klasifikasi Sentimen	8
BAB III PENUTUP.....	10
3.1. Kesimpulan	10

BAB I

PENDAHULUAN

1.1. Latar Belakang

Perkembangan teknologi digital dan meningkatnya penggunaan perangkat mobile menjadikan layanan streaming video berbasis aplikasi sebagai salah satu sarana hiburan yang banyak digunakan oleh masyarakat. Di Indonesia, Netflix dan Vidio merupakan dua layanan streaming yang populer dengan karakteristik konten dan layanan yang berbeda. Netflix dikenal dengan penyediaan film dan serial internasional, sedangkan Vidio lebih berfokus pada konten lokal, siaran televisi, dan acara olahraga. Perbedaan karakteristik tersebut menjadi alasan dipilihnya kedua layanan ini sebagai objek analisis.

Sumber data teks dalam laporan ini berasal dari ulasan pengguna aplikasi Netflix dan Vidio. Ulasan aplikasi dipilih karena memuat pendapat, pengalaman, serta penilaian pengguna secara langsung terhadap layanan yang digunakan. Melalui ulasan tersebut, pengguna menyampaikan tingkat kepuasan maupun ketidakpuasan terhadap berbagai aspek layanan, seperti kualitas konten, performa aplikasi, dan kemudahan penggunaan. Dengan demikian, ulasan aplikasi dapat dimanfaatkan untuk memahami persepsi dan sentimen pengguna.

Fenomena yang ingin dilihat melalui analisis ini adalah adanya perbedaan kecenderungan sentimen pengguna terhadap dua layanan streaming yang memiliki segmentasi dan karakteristik berbeda. Meskipun Netflix dan Vidio sama-sama menyediakan layanan streaming video, tanggapan pengguna terhadap kualitas layanan dan konten yang ditawarkan dapat berbeda, yang tercermin dari sentimen positif, netral, negatif pada ulasan pengguna.

Untuk melihat fenomena tersebut, dilakukan analisis sentimen komparatif terhadap ulasan pengguna Netflix dan Vidio. Dengan memanfaatkan teknik text mining dan machine learning, data teks ulasan diolah untuk mengidentifikasi kecenderungan sentimen serta pola bahasa yang dominan. Hasil analisis ini diharapkan dapat memberikan gambaran mengenai perbedaan persepsi pengguna terhadap kedua layanan streaming serta menjadi bahan pertimbangan dalam peningkatan kualitas layanan.

1.2. Tujuan

Adapun tujuan dari analisis sentimen komparatif pengguna layanan streaming Netflix dan Vidio adalah sebagai berikut:

1. Mengetahui kecenderungan sentimen pengguna terhadap aplikasi Netflix dan Vidio berdasarkan ulasan yang diberikan oleh pengguna.

2. Mengidentifikasi kata dan frasa dominan yang sering muncul pada ulasan masing-masing aplikasi.
3. Mengidentifikasi polaritas sentimen ulasan pengguna yang terbagi ke dalam kategori positif, negatif, dan netral menggunakan pendekatan berbasis leksikon.
4. Mengklasifikasikan sentimen ulasan pengguna Netflix dan Vidio menggunakan metode machine learning (Naive Bayes, SVM, Logistic Regression) dan deep learning (LSTM) untuk membandingkan performa model.
5. Membandingkan persepsi pengguna terhadap kualitas layanan, performa aplikasi, dan konten berdasarkan hasil analisis teks.

BAB II

HASIL PEMBAHASAN

2.1. Akuisisi Data

Data yang digunakan dalam analisis ini berasal dari ulasan pengguna aplikasi Netflix dan Vidio yang tersedia di Google Play Store. Aplikasi ini dipilih karena menyediakan ulasan pengguna yang mencerminkan persepsi nyata terhadap kualitas layanan, performa aplikasi, dan konten yang disediakan. Selain itu, ulasan bersifat terbuka dan mudah diakses, sehingga sesuai sebagai sumber data teks untuk analisis sentimen.

Proses pengambilan data dilakukan menggunakan teknik web scraping dengan memanfaatkan library Google Play Scraper pada bahasa pemrograman Python. Library tambahan seperti pandas digunakan untuk menyimpan dan mengelola data dalam bentuk tabel, sedangkan numpy digunakan untuk manipulasi data. Data diambil secara otomatis dari halaman aplikasi Netflix dan Vidio, kemudian disimpan dalam format CSV untuk tahap pra-pemrosesan dan analisis. Pengambilan data dilakukan secara publik tanpa melibatkan akun pengguna tertentu.

Data ulasan dikumpulkan dengan parameter bahasa Indonesia dan wilayah Indonesia, sehingga merepresentasikan pendapat pengguna di Indonesia. Proses pengambilan data dilakukan pada 27 Desember 2025 melalui Google Play Store, dengan fokus pada ulasan terbaru dari pengguna aplikasi Netflix dan Vidio. Setiap aplikasi dikumpulkan sebanyak 3.000 ulasan, sehingga total dataset berjumlah 6.000 ulasan. Setiap data ulasan memiliki tiga atribut utama:

1. Review – teks ulasan pengguna, digunakan sebagai data utama untuk pra-pemrosesan, analisis sentimen, dan analisis kata/ frasa dominan.
2. Rating – penilaian aplikasi dalam skala 1–5, digunakan sebagai informasi pendukung.
3. App – nama aplikasi (Netflix atau Vidio), digunakan untuk analisis komparatif antar aplikasi.

2.2. Pra-pemrosesan Data

Pra-pemrosesan data dilakukan untuk membersihkan teks ulasan sebelum dianalisis. Tahap pertama adalah case folding, yaitu mengubah seluruh teks menjadi huruf kecil untuk menyeragamkan format penulisan. Selanjutnya dilakukan cleaning dengan menghapus URL, mention, hashtag, angka, tanda baca, spasi berlebih, serta emoji, sehingga teks terbebas dari karakter yang tidak relevan. Setelah teks bersih, dilakukan tokenizing untuk memecah kalimat menjadi kata-kata, kemudian filtering melalui penghapusan stopword menggunakan daftar stopword dari NLTK dan Sastrawi guna menghilangkan kata umum yang tidak berpengaruh dalam analisis. Tahap terakhir adalah

stemming menggunakan Stemmer Sastrawi untuk mengubah kata ke bentuk dasarnya, sehingga diperoleh teks yang lebih terstandarisasi dan siap digunakan pada proses analisis selanjutnya.

2.3. Metodologi Analisis

2.3.1. Analisis Deskriptif

Analisis deskriptif dilakukan untuk memahami karakteristik umum data teks ulasan pengguna pada aplikasi Netflix dan Vidio. Pada tahap ini digunakan visualisasi Word Cloud untuk menampilkan kata-kata yang paling sering muncul dalam ulasan, sehingga memberikan gambaran awal mengenai topik dan fokus utama opini pengguna pada masing-masing aplikasi.

Selain itu, diterapkan analisis N-Gram, yaitu bigram dan trigram, untuk mengidentifikasi kombinasi dua dan tiga kata yang sering muncul secara berurutan dalam teks ulasan. Pendekatan ini bertujuan menangkap konteks penggunaan kata secara lebih baik serta mengetahui pola bahasa yang khas dalam opini pengguna. Frekuensi kemunculan bigram dan trigram kemudian dihitung untuk menentukan frasa dominan yang sering digunakan dalam ulasan Netflix dan Vidio.

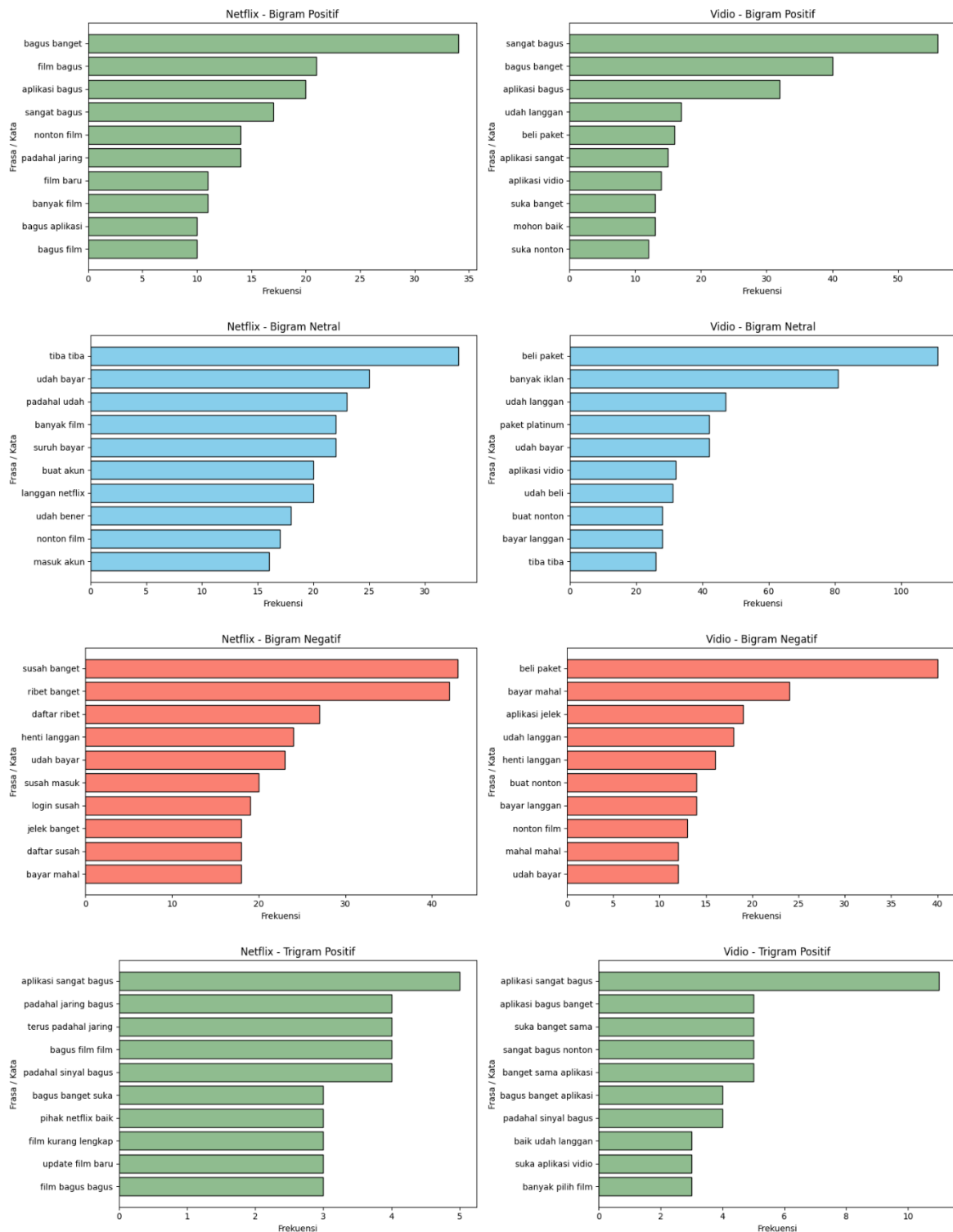
2.3.2. Analisis Sentimen

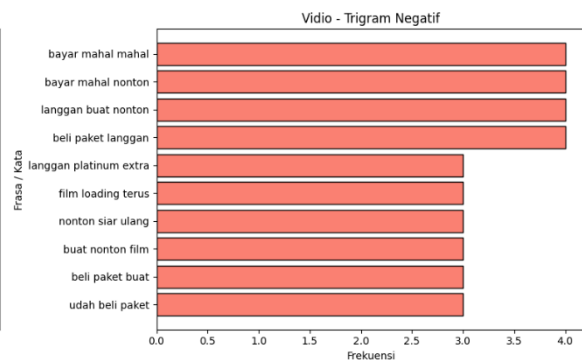
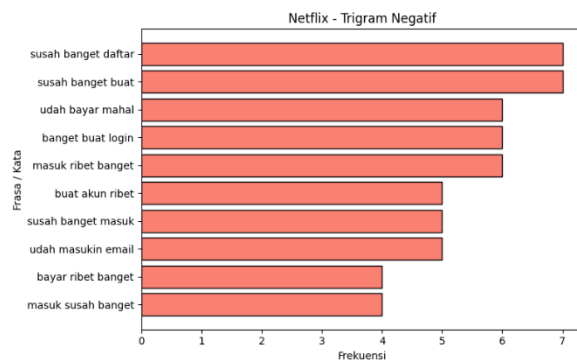
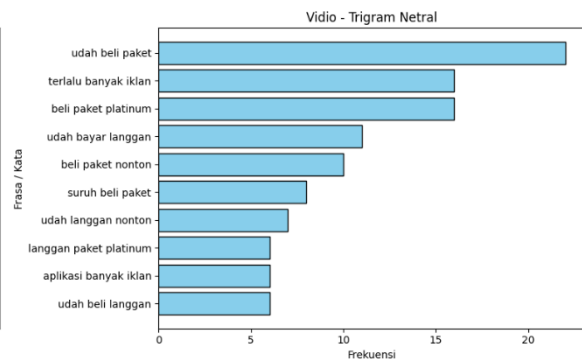
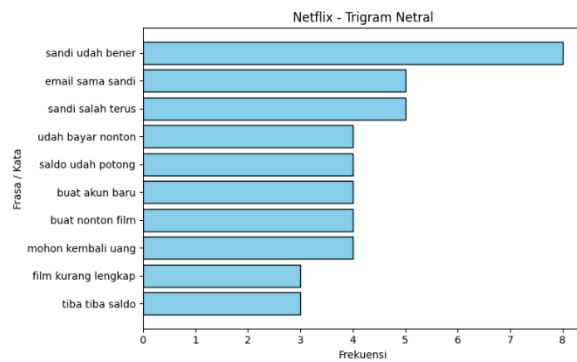
Analisis sentimen digunakan untuk mengelompokkan ulasan pengguna ke dalam kategori positif, negatif, dan netral. Metode yang digunakan adalah pendekatan berbasis lexicon, yaitu mencocokkan kata-kata hasil pra-pemrosesan dengan kamus kata positif dan negatif. Setiap ulasan diberikan skor berdasarkan jumlah kemunculan kata positif dan negatif. Ulasan dengan skor positif lebih tinggi dikategorikan sebagai sentimen positif, skor negatif lebih tinggi sebagai sentimen negatif, sedangkan skor yang seimbang dikategorikan sebagai sentimen netral.

2.4. Hasil

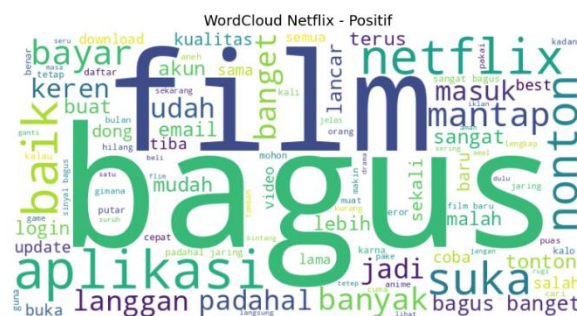
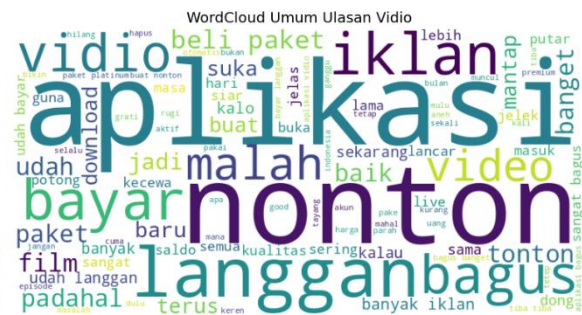
Hasil analisis sentimen terhadap ulasan pengguna Netflix dan Vidio disajikan dalam bentuk visualisasi komparatif untuk membandingkan kecenderungan opini pengguna pada kedua layanan streaming. Visualisasi yang digunakan meliputi grafik distribusi frekuensi kata/frasa (bigram) dan Word Cloud yang ditampilkan berdampingan antara kedua aplikasi.

Grafik distribusi sentimen menampilkan frasa yang paling sering muncul pada kategori sentimen positif, negatif, dan netral di masing-masing aplikasi, sehingga terlihat pola bahasa yang dominan dalam ulasan pengguna Netflix dan Vidio





Word Cloud umum menunjukkan kata-kata yang paling sering muncul dari seluruh ulasan, sedangkan Word Cloud berdasarkan kategori sentimen menampilkan kata-kata dominan pada sentimen positif, negatif, dan netral. Visualisasi ini membantu melihat bentuk apresiasi, keluhan, serta perbedaan fokus opini pengguna antara Netflix dan Vidio secara lebih jelas.





Interpretasi hasil pada grafik menunjukkan bahwa Netflix memiliki proporsi sentimen positif yang lebih dominan dibandingkan sentimen negatif, yang mengindikasikan sebagian besar pengguna merasa puas terhadap kualitas layanan dan konten yang disediakan. Sebaliknya, Vidio menunjukkan jumlah sentimen negatif yang relatif lebih tinggi dibandingkan Netflix, yang menandakan masih adanya keluhan pengguna terkait performa aplikasi maupun layanan teknis. Selain itu, Word Cloud pada Netflix, kata yang sering muncul berkaitan dengan kualitas konten, film, serial, dan pengalaman menonton, sehingga mencerminkan fokus utama pengguna adalah pada variasi dan mutu tayangan. Sementara itu, pada Vidio, kata-kata dominan lebih banyak berkaitan dengan aplikasi, error, buffering, dan langganan, yang menunjukkan bahwa pengguna lebih sering membahas aspek teknis dan performa layanan.

Analisis menunjukkan beberapa pola menarik dalam ulasan pengguna. Pada Netflix, meskipun mayoritas ulasan bersentimen positif karena kualitas konten yang baik, masih ditemukan keluhan terkait harga langganan dan aturan akun. Sementara itu, pada Vidio, kata-kata seperti error, buffering, dan lag sering muncul, bahkan pada ulasan netral, yang menandakan bahwa masalah teknis cukup sering dialami pengguna. Selain itu, kata-kata seperti gratis dan promo muncul pada ulasan positif Vidio, menunjukkan bahwa faktor harga menjadi daya tarik bagi pengguna. Temuan ini memperlihatkan bahwa Netflix lebih banyak diapresiasi dari sisi konten, sedangkan Vidio lebih banyak disorot dari sisi performa aplikasi.

2.5. Pemodelan Klasifikasi Sentimen

Tahap permodelan ini menunjukkan bahwa analisis sentimen tidak hanya bersifat deskriptif, tetapi juga dapat diterapkan secara prediktif melalui pendekatan *machine learning* yang terukur. Tahap ini merupakan proses klasifikasi sentimen dengan menerapkan algoritma *machine learning* dan *deep learning* pada ulasan pengguna aplikasi Netflix dan Vidio. Teks ulasan yang telah melalui tahap pra-pemrosesan kemudian ditransformasikan ke dalam bentuk representasi numerik menggunakan metode Term Frequency–Inverse Document Frequency (TF-IDF). Metode ini memberikan bobot pada kata-kata penting yang sering muncul dalam dokumen, sehingga mampu merepresentasikan karakteristik teks secara lebih informatif dan mendukung model dalam mengenali pola sentimen secara lebih akurat. Dataset masing-masing aplikasi dibagi menjadi data latih dan data uji dengan perbandingan 80% data latih dan 20% data uji.

Model klasifikasi pertama yang digunakan adalah Naive Bayes. Pada data ulasan Netflix, model Naive Bayes menghasilkan akurasi sebesar 77%, sedangkan pada data ulasan Vidio diperoleh akurasi sebesar 74%. Hasil ini menunjukkan bahwa model Naive Bayes mampu melakukan klasifikasi sentimen dengan tingkat ketepatan tertentu dan digunakan sebagai model pembanding awal.

Model klasifikasi kedua yang digunakan adalah Support Vector Machine (SVM). Hasil evaluasi menunjukkan bahwa model SVM menghasilkan akurasi sebesar 87% pada Netflix dan 89% pada Vidio, yang lebih tinggi dibandingkan model Naive Bayes. Hal ini menunjukkan bahwa SVM memiliki kemampuan yang lebih baik dalam memisahkan kelas sentimen pada data teks berbasis TF-IDF.

Model klasifikasi ketiga yang digunakan adalah Logistic Regression. Model ini menghasilkan akurasi sebesar 83% pada Netflix dan 81% pada Vidio, dengan performa yang relatif stabil dan kompetitif. Logistic Regression mampu memodelkan hubungan linier antar fitur TF-IDF tanpa asumsi independensi kata seperti pada Naive Bayes.

Untuk pendekatan deep learning, digunakan model Long Short-Term Memory (LSTM). Berdasarkan hasil pengujian, model LSTM memperoleh akurasi tertinggi, yaitu sebesar 90% pada Netflix dan 91% pada Vidio. Keunggulan model LSTM terletak pada kemampuannya dalam menangkap hubungan sekuensial dan konteks antar kata dalam teks ulasan, sehingga menghasilkan prediksi sentimen yang lebih akurat.

Perbedaan nilai akurasi klasifikasi sentimen antara aplikasi Netflix dan Vidio dipengaruhi oleh karakteristik data ulasan masing-masing aplikasi. Ulasan Netflix cenderung memiliki bahasa yang lebih konsisten dan deskriptif dengan distribusi sentimen yang relatif seimbang, sehingga pola kata dapat direpresentasikan dengan baik oleh metode TF-IDF dan dipelajari secara efektif oleh model klasifikasi. Sebaliknya, ulasan Vidio lebih bervariasi, sering bersifat singkat, informal, serta banyak mengandung keluhan teknis, yang menyebabkan kompleksitas pola teks lebih tinggi. Kondisi ini memengaruhi kinerja model, khususnya pada algoritma berbasis fitur linier. Namun, model LSTM mampu memberikan akurasi yang lebih tinggi dan stabil pada kedua aplikasi karena kemampuannya dalam menangkap hubungan sekuensial dan konteks antar kata dalam teks ulasan.

Berdasarkan hasil pemodelan tersebut, dapat disimpulkan bahwa model LSTM memberikan kinerja terbaik dalam klasifikasi sentimen ulasan aplikasi Netflix dan Vidio, diikuti oleh Support Vector Machine, Logistic Regression, dan Naive Bayes. Tahap pemodelan ini menunjukkan bahwa analisis sentimen dapat diterapkan secara prediktif menggunakan pendekatan machine learning yang terukur dan dapat dievaluasi secara kuantitatif.

BAB III

PENUTUP

3.1. Kesimpulan

Berdasarkan hasil analisis sentimen terhadap ulasan pengguna aplikasi Netflix dan Vidio, dapat disimpulkan bahwa kedua layanan streaming tersebut memiliki karakteristik sentimen yang berbeda. Netflix cenderung memperoleh sentimen positif yang lebih dominan, yang menunjukkan tingkat kepuasan pengguna yang lebih tinggi terhadap kualitas konten dan pengalaman menonton. Sementara itu, Vidio memiliki proporsi sentimen negatif yang relatif lebih besar, yang banyak dipengaruhi oleh keluhan pengguna terkait performa aplikasi dan masalah teknis layanan.

Hasil analisis kata dan frasa dominan menunjukkan adanya perbedaan fokus pembahasan pada masing-masing aplikasi, di mana ulasan Netflix lebih banyak menyoroti kualitas konten dan fitur, sedangkan ulasan Vidio didominasi oleh keluhan teknis dan stabilitas aplikasi. Hal ini memperkuat perbedaan persepsi pengguna terhadap kedua layanan.

Hasil pemodelan klasifikasi sentimen menggunakan empat algoritma menunjukkan bahwa seluruh model mampu melakukan klasifikasi dengan tingkat akurasi yang berbeda. Naive Bayes memberikan performa terendah namun tetap dapat digunakan sebagai model dasar, sedangkan Support Vector Machine dan Logistic Regression menunjukkan peningkatan akurasi yang signifikan. Model Long Short-Term Memory (LSTM) menghasilkan akurasi tertinggi pada kedua aplikasi, yang menandakan kemampuannya dalam memahami konteks dan urutan kata pada teks ulasan secara lebih baik.

Secara keseluruhan, analisis ini menunjukkan bahwa pendekatan text mining dan machine learning dapat dimanfaatkan secara efektif untuk memahami persepsi pengguna terhadap layanan streaming. Hasil yang diperoleh dapat digunakan sebagai bahan evaluasi bagi penyedia layanan dalam meningkatkan kualitas konten, performa aplikasi, serta kepuasan pengguna.