

PENERAPAN ALGORITMA NAÏVE BAYE UNTUK ANALISIS SENTIMEN APLIKASI SPOTIFY PADA GOOGLE PLAY STORE

Novrian^{1*}, Hari Soetanto²

^{1,2}Teknik Informatika, Fakultas Teknologi Informasi, Universitas Budi Luhur, Jakarta, Indonesia

Email: ^{1*}2111500621@student.budiluhur.ac.id, ²hari.soetanto@budiluhur.ac.id
(* : corresponding author)

Abstrak- Perkembangan pesat layanan streaming musik seperti *Spotify* telah mengubah cara pengguna menikmati konten audio. Meskipun popularitasnya tinggi, aplikasi ini sering menerima ulasan beragam dari pengguna yang dapat memengaruhi citra dan kualitas layanan. Mengingat jumlah ulasan yang besar, analisis manual menjadi tidak efisien. Oleh karena itu, penelitian ini bertujuan untuk mengatasi masalah tersebut dengan menerapkan algoritma *Naïve Bayes* untuk menganalisis sentimen ulasan pengguna *Spotify*. Data ulasan dikumpulkan dari *Google Play Store* dan melalui tahap pra-pemrosesan seperti *case folding*, *cleansing*, *tokenization*, dan *stemming* untuk memastikan kualitas data. Model klasifikasi sentimen dibangun menggunakan algoritma *Naïve Bayes*. Hasil pengujian menunjukkan bahwa model ini efektif dalam mengklasifikasikan ulasan dengan akurasi sebesar 73,79%. Implementasi *Naïve Bayes* berhasil mengidentifikasi sentimen positif dan negatif, meskipun masih terdapat tantangan pada ulasan dengan konteks ambigu. Hasil analisis ini diharapkan dapat memberikan wawasan berharga bagi pengembang *Spotify* untuk meningkatkan kualitas layanan dan kepuasan pengguna.

Kata Kunci: Analisis Sentimen, Ulasan Pengguna, Naive Bayes, Google Play Store, Spotify

APPLICATION OF THE NAÏVE BAYES ALGORITHM FOR SENTIMENT ANALYSIS OF THE SPOTIFY APPLICATION ON THE GOOGLE PLAY STORE

Abstract- The rapid development of music streaming services such as *Spotify* has changed the way users consume audio content. Despite its popularity, the application receives a wide range of user reviews that can impact its brand image and service quality. Given the large volume of reviews, manual analysis is highly inefficient. Therefore, this study aims to address this problem by applying the *Naïve Bayes* algorithm to perform sentiment analysis on *Spotify* user reviews. Review data were collected from the *Google Play Store* and went through a pre-processing phase, including *case folding*, *cleansing*, *tokenization*, and *stemming*, to ensure data quality. A sentiment classification model was then built using the *Naïve Bayes* algorithm. The test results show that this model is effective in classifying reviews, achieving an accuracy of 77.24%. The implementation of *Naïve Bayes* successfully identified positive and negative sentiments, although challenges remain for reviews with ambiguous contexts. The findings of this analysis are expected to provide valuable insights for *Spotify* developers to improve service quality and user satisfaction.

Keywords: Sentiment Analysis, User Reviews, Naïve Bayes, Google Play Store, Spotify

1. PENDAHULUAN

Perkembangan teknologi internet dan perangkat seluler telah berkontribusi secara signifikan terhadap perubahan dalam industri musik serta memicu kemunculan layanan *streaming* musik. Layanan *streaming* musik kini menjadi metode yang paling diminati bagi para pengguna untuk menikmati lagu-lagu. Melalui layanan ini, para pengguna bisa menikmati akses tanpa batas ke berbagai pilihan katalog musik [1]. *Spotify* sebagai salah satu aplikasi *streaming* yang memungkinkan pengguna untuk mendengarkan musik, *podcast*, atau pun konten audio lainnya dari seluruh dunia secara *online*. *Spotify* merupakan salah satu platform musik *streaming* yang banyak diminati dan telah direkomendasikan oleh para editor di *Google Play*. Aplikasi ini telah diunduh lebih dari satu juta kali dan mendapatkan rata-rata penilaian 4.4. Dengan jumlah pengguna yang tinggi, sangat penting untuk memahami keinginan dan ekspektasi mereka untuk meningkatkan kualitas serta daya saing dari aplikasi [2]. Sebelum hadirnya platform *streaming* musik seperti *Spotify*, orang-orang umumnya menikmati musik lewat radio, membeli CD atau kaset, atau bahkan mengunduh lagu secara tidak sah melalui internet [3]. Berdasarkan penelitian mengenai penetrasi internet di Indonesia tahun 2024 yang diselenggarakan oleh APJII, kelompok usia 28-43 tahun mendominasi jumlah pengguna internet, mencakup 93,17% dari total 221 juta pengguna di Indonesia. Dari angka tersebut, 56,07% menikmati platform *streaming* musik menggunakan *handphone* atau *tablet*, dengan tingkat pemakaian perangkat ini mencapai 89,44%, berbeda dengan akses melalui aplikasi TV berbasis internet yang

hanya mencapai 11,02% [4]. Platform pemutar audio seperti *Spotify*, *Apple Music*, dan *YouTube Music* mengalami pertumbuhan signifikan dalam penawaran mereka dan menyuguhkan berbagai inovasi yang menarik. Di kuartal kedua tahun 2021, *Spotify* berhasil merebut 31% dari total pangsa pasar dalam industri pemutaran musik global [5].

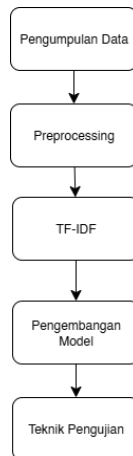
Analisis sentimen adalah sebuah proses untuk menentukan sentimen atau opini dari seseorang yang diwujudkan dalam bentuk teks dan bisa dikategorikan sebagai sentimen positif atau negatif [6]. Metode yang diterapkan pada analisis sentimen ini yaitu algoritma *Naïve Bayes* algoritma ini efektif untuk pengelompokan teks secara akurasi yang tinggi dan memadai untuk teks yang bersifat besar dan kompleks [7]. Algoritma *Naïve Bayes* adalah salah satu Pendekatan pengelompokan berbasis probabilitas yang didasarkan pada *Teorema Bayes*. Algoritma ini menggunakan anggapan dasar bahwa setiap atribut dalam kumpulan data terpisah satu sama lain sering kali disebut "*naïve*" atau sederhana. Penggunaan metode algoritma *Naïve Bayes* diharapkan dapat memberikan kontribusi pada literatur terkait analisis sentimen dan memahami potensi metode dalam mengatasi sentimen pengguna, diperkuat dalam konteks aplikasi layanan *streaming* musik *online* [8]. Pada penelitian ini, analisis sentimen akan digunakan untuk menentukan tingkat akurasi berdasarkan data komentar atau pendapat yang didapati di media sosial *Twitter* atau aplikasi *Google Play Store*.

Analisis sentimen adalah proses mengidentifikasi dan mengklasifikasikan opini pengguna dari teks, yang bisa dikategorikan positif, negatif, atau netral. Algoritma *Naïve Bayes* sering digunakan karena efisien dan akurat untuk teks besar dan kompleks. Penelitian sebelumnya telah menggunakan *Naïve Bayes* untuk analisis sentimen, seperti pada aplikasi perbankan digital *Bank Saqu*. Studi ini menunjukkan *Naïve Bayes* efektif dalam mengklasifikasikan opini pengguna [9].

Untuk menyempurnakan penelitian-penelitian yang sudah ada, Penelitian ini tertarik untuk membahas topik mengenai penerapan algoritma *Naïve Bayes* untuk analisis sentimen ulasan pengguna pada aplikasi *Spotify*, dapat memberikan wawasan yang lebih cepat dalam mengidentifikasi pola-pola penting dalam ulasan pelanggan dan mengambil tindakan yang tepat untuk meningkatkan kualitas produk serta layanan mereka.

2. METODE PENELITIAN

Penelitian ini menggunakan metode *Naïve Bayes* dalam menganalisis sentimen ulasan pengguna aplikasi *Spotify*. Proses penelitian ini dirancang secara sistematis melalui serangkaian tahapan yang dimulai dari pengumpulan data, *preprocessing*, pembobotan kata dengan TF-IDF, pengembangan model, hingga pengujian model.



Gambar 1. Penerapan Metode *Naïve Bayes*

2.1 Metode Pengumpulan Data

Tahapan paling awal yang dilakukan adalah mengumpulkan data untuk langkah proses pengujian. Data yang digunakan dalam penelitian ini berasal dari komentar aplikasi *Spotify* yang terdapat di *Google Play Store*, yang diperoleh melalui teknik pengambilan data dengan cara *scrapping* menggunakan *Google Collab* serta *Library Google Play Scraper*. Data diperoleh dari tanggal 4 Juni 2025 hingga 6 Juni 2025 dengan total data 483 yang kemudian data tersebut disimpan dalam format *.xlsx*.

Tabel 1. Sampel Data Ulasan Aplikasi *Spotify*

No	Username	Ulasan
1.	Rfqy Ikyy	Lengkap
2.	Safira Riselly	Makin kikir, shuffle nya ga ngotak
3.	Nayaka Khairy Aydin Palgunadi	Mantap
4.	Yanti Trisnawati	Sangat bagus, tetapi di saya tidak ada liriknya
5.	Mas Don	Bagus banget
...
484.	Suratman Adi	Bagus

2.2 Tahapan *Preprocessing*

Preprocessing dalam penambangan teks merupakan langkah awal dalam pengolahan data teks untuk mengonversinya menjadi format yang lebih sistematis dan dapat dianalisis oleh algoritma pembelajaran mesin. Proses ini merupakan langkah yang paling krusial dalam analisis sentimen, sebab semakin terorganisir data yang dimanfaatkan, maka hasil prediksi yang diperoleh akan semakin tepat [10]. Tujuan dari proses pengolahan data ini adalah untuk mempermudah analisis data di tahap berikutnya. Pada tahap ini, dilakukan pembersihan serta persiapan data supaya data yang akan dianalisis menjadi lebih terstruktur dan siap untuk penelitian. Teks yang diperoleh dari beragam sumber, seperti artikel, ulasan, atau platform media sosial, umumnya memiliki elemen yang tidak relevan atau berlebihan [11].

Setelah data dikumpulkan, langkah selanjutnya adalah *preprocessing*. Tujuan dari tahap *preprocessing* ini adalah untuk membersihkan *dataset* yang diambil dari ulasan *Google Play Store*, sehingga data tersebut menjadi lebih terstruktur dan siap untuk analisis lebih lanjut. Proses ini menjamin bahwa data untuk pelabelan dan pengujian klasifikasi sentimen berada dalam kondisi yang optimal, sehingga hasil yang diperoleh dapat mencapai tingkat akurasi yang tinggi. *Preprocessing* melibatkan berbagai langkah, seperti menghapus karakter yang tidak diperlukan, merapikan teks, dan menyaring informasi yang relevan untuk Analisis Sentimen. Berikut adalah langkah-langkah yang dilakukan dalam tahapan *preprocessing*:

a. *Case Folding*

Case Folding adalah tahapan untuk mengubah keseluruhan huruf menjadi huruf kecil (*lowercase*). Bertujuan untuk menyamakan format teks agar lebih konsisten. Misalnya kata ‘Keren’ atau ‘KEREN’ akan diubah menjadi ‘keren’.

b. *Cleansing*

Cleansing data adalah proses penting dalam tahap *preprocessing* yang bertujuan untuk membersihkan teks dari elemen-elemen yang tidak relevan, sehingga data menjadi lebih terstruktur dan siap untuk analisis. Tahapan ini melakukan pembersihan pada teks seperti emoji, simbol, url, dan hastag. Contohnya, kata ‘#spotify’ menjadi ‘spotify’, ‘bagus!’ menjadi ‘bagus’ dan ‘https://open.spotify.com/user/spotify’ akan dihapus dari kalimat.

c. *Tokenization*

Tokenization adalah proses analisis teks di mana deretan karakter dalam sebuah kalimat dipecah menjadi unit-unit yang lebih kecil yang dikenal sebagai token. *Tokenization* bisa berupa istilah, ungkapan, tanda baca, atau simbol tunggal. Dengan adanya tokenisasi, mesin dapat menganalisis teks dengan lebih terorganisir karena isi teks telah dibagi menjadi elemen-elemen dasar. *Tokenization* adalah langkah dalam memecah teks menjadi elemen-elemen kata. Sebagai contoh: ‘bagus sekali’ setelah melalui proses ini akan terpisah menjadi dua kata yaitu: ‘bagus’ dan ‘sekali’.

d. Normalisasi

Normalisasi adalah proses penting dalam *preprocessing* data teks yang bertujuan untuk mengubah kata-kata tidak baku atau tidak standar menjadi bentuk baku atau standar. Tahapan ini untuk memastikan keseragaman dalam data teks, karena seringkali dalam percakapan sehari-hari atau tulisan informal, orang menggunakan berbagai variasi kata yang sebenarnya memiliki makna sama. Contohnya, kata-kata seperti ‘bgt’ menjadi ‘banget’, ‘yg’ menjadi ‘yang’, lalu ‘dgn’ menjadi ‘dengan’, dan ‘sdh’ menjadi ‘sudah’

e. *Stopword Removal*

Tahap ini bertujuan untuk menyaring dan menghilangkan kata-kata umum yang sering muncul namun tidak memiliki kontribusi signifikan terhadap makna inti dari sebuah teks. Kata-kata ini dikenal sebagai *stopwords*. Contoh *stopwords* dalam Bahasa Indonesia antara lain “yang”, “di”, “dan”, “adalah”, “dengan”, “ke”, “ini”, dan sebagainya. Dengan menghapus kata-kata tersebut, model dapat lebih fokus pada kata-kata kunci yang relevan, sehingga mengurangi noise dan meningkatkan akurasi analisis.

f. *Stemming*

Tahap terakhir dalam pra-pemrosesan teks adalah *Stemming*. *Stemming* adalah proses untuk mengubah setiap kata ke dalam bentuk dasarnya (kata dasar atau stem) dengan menghilangkan imbuhan (awalan, sisipan, akhiran). Tujuannya adalah untuk mengelompokkan berbagai variasi kata yang memiliki akar makna yang sama menjadi satu representasi tunggal. Sebagai contoh, kata "menganalisis", "penganalisan", dan "dianalisis" semuanya akan dikurangi menjadi kata dasar yang sama, yaitu "analisis". Proses ini sangat penting agar algoritma tidak memperlakukan kata-kata tersebut sebagai fitur yang berbeda, sehingga generalisasi model menjadi lebih efektif.

2.3 Pembobotan Kata Menggunakan TF-IDF

TF-IDF adalah singkatan dari Term Frequency-Inverse Document Frequency, sebuah teknik statistik yang populer untuk digunakan dalam bidang penambangan teks. Tujuannya adalah untuk mengukur seberapa penting sebuah kata atau frasa dalam sebuah dokumen di dalam suatu koleksi dokumen. Timbangan TF-IDF adalah perhitungan statistik untuk mengetahui seberapa penting sebuah kata bagi sebuah dokumen dalam koleksi dokumen [12]. Istilah bisa muncul sebagai kata, ungkapan, atau elemen pengindeksan lain yang menghasilkan dokumen yang dipakai untuk mengidentifikasi konteks dari dokumen tersebut. Istilah bisa muncul sebagai kata, ungkapan, atau elemen pengindeksan lain yang menghasilkan dokumen yang dipakai untuk mengidentifikasi konteks dari dokumen tersebut. Selanjutnya, setiap kata akan diberikan tanda, yaitu nilai bobot istilah [13].

Berikut adalah rumus perhitungan TF-IDF :

a. Rumus perhitungan TF

$$Tf(t, d) = \frac{\text{Jumlah kemunculan kata } t \text{ dalam dokumen } d}{\text{Jumlah total kata dalam dokumen } d} \quad (2.1)$$

b. Rumus perhitungan IDF

$$IDF(t) = \log \frac{N}{1+df} \quad (2.2)$$

c. Rumus perhitungan TF-IDF :

$$TF-IDF = TF(t,d)*IDF(t) \quad (2.3)$$

2.4 Tahapan Pengembangan Model

Tahapan pengembangan model yaitu pengujian data latih dan data uji menggunakan algoritma *Naive Bayes* dengan dataset yang telah menjalani tahap *text-preprocessing* dan pembobotan kata menggunakan TF-IDF. Selanjutnya, Dataset dibagi menjadi dua bagian, yaitu data pelatihan dan data pengujian yang akan digunakan untuk analisis.

2.5 Teknik Pengujian

Setelah melalui tahapan pengembangan model, tahapan teknik pengujian adalah proses untuk mengevaluasi kinerja model menggunakan data uji yang belum pernah dilihat model sebelumnya. Pada tahap ini, akan terlihat metrik seperti akurasi, presisi, *recall*, dan *f1-score*, serta menganalisis *confusion matrix*.

3. HASIL DAN PEMBAHASAN

Bab ini menyajikan hasil dari setiap tahapan penelitian yang telah dijelaskan pada bab sebelumnya, dimulai dari pengumpulan data hingga evaluasi model. Pembahasan akan berfokus pada hasil yang diperoleh dan interpretasinya, serta bagaimana setiap tahapan berkontribusi pada pencapaian tujuan penelitian.

3.1 Pengumpulan Data

Pada tahapan "Pengumpulan Data", sistem ini memungkinkan pengguna mengunggah file dataset, lalu secara otomatis menampilkan ringkasan jumlah data positif dan negatif. Selain itu, sistem juga menyajikan pratinjau data misalnya (kolom ulasan dan label) untuk memverifikasi data yang diunggah. Tahap pengumpulan data dapat dilihat pada tabel 2.

Tabel 2. Pengumpulan Data

Ulasan	Label
Lengkap	POSITIVE
Makin kikir, shuffle nya ga ngotak	POSITIVE
Mantap	NEGATIVE
Sangat bagus, tetapi di saya tidak ada liriknya	POSITIVE
Masih banyak yang kurang	NEGATIVE
.....
Bagus banget	POSITIVE

3.2 Preprocessing

Setelah tahapan pengumpulan data, langkah berikutnya adalah melakukan preprocessing. Sesuai dengan penjelasan dari sub bab (2.2), urutan dari tahapan ini terdiri dari 7 proses utama yaitu: *case folding*, *cleansing*, normalisasi, *tokenization*, normalisasi, *stopword removal*, *stemming*.

a. Case Folding

Pada tahap *case folding* ini digunakan untuk menyeragamkan kata yang berkomposisi huruf kapital atau huruf besar menjadi huruf kecil (*lowercase*). Hasil *case folding* tersebut dapat dilihat pada tabel 3.

Tabel 3. Hasil Case Folding

Sebelum	Sesudah
Bagus tapi kurang di saat lagu yang di play malah lagu lain yang muncul	bagus tapi kurang di saat lagu yang di play malah lagu lain yang muncul
BAGUS BANGET BISA BUAT MAIN SAMBIL DENERIN LAGU ENAK BANGET	bagus banget bisa buat main sambil dengerin lagu enak banget

b. Cleansing

Tahapan dilanjutkan dengan *cleansing*, teks yang tidak perlu akan disaring dan dihilangkan. Elemen-elemen yang tidak relevan seperti tanda baca, angka, atau simbol dihapus dari teks. Hasil *cleansing* tersebut dapat dilihat pada tabel 4.

Tabel 4. Hasil Cleansing

Sebelum	Sesudah
makin kikir, shuffle nya tidak ngotak	makin kikir shuffle nya tidak ngotak
plis iklan yang "aku cinta kamu ihh !!!" Hapus dong dengerinnya jijik tau gue yang suka KPOP aja kagak Sampek gitu plis malu malu in Spotify deh	plis iklan yang aku cinta kamu ihh hapus dong dengerinnya jijik tau saya yang suka kpop saja kagak sampek gitu plis malu malu in Spotify deh

c. Tokenization

Setelah itu dilanjutkan tahapan *cleansing*, tahapan mendasar di mana kumpulan teks yang besar diuraikan atau dipisahkan menjadi bagian-bagian yang lebih kecil dan terpisah. Bagian-bagian ini, yang disebut 'token', seringkali merupakan kata-kata tunggal, namun bisa juga mencakup potongan-potongan lain yang dianggap sebagai unit penting. Hasil *tokenization* tersebut dapat dilihat pada tabel 5.

Tabel 5. Hasil Tokenization

Sebelum	Sesudah
sangat bagus, tetapi di Saya tidak ada lirik nya	sangat, bagus, tetapi, di, saya, tidak, ada, lirik, nya
aplikasi ini dapat dimatikan tapi juga bisa di dengar kan oleh kita/dimasukkan ke sound yang kalian punya	aplikasi, ini, dapat, dimatikan, tapi, juga, bisa, di, dengar, kan, oleh, kita,

dimasukkan, ke, sound, yang, kalian,
punya

d. Normalisasi

Setelah melalau tahapan tokenization, selanjutnya adalah tahapan normalisasi, di mana kata-kata tidak baku atau tidak standar diubah menjadi bentuk baku. Hasil normalisasi ini penting untuk menyeragamkan data teks, memastikan konsistensi, dan mempersiapkan teks. Hasil normalisasi dapat dilihat pada tabel 6

Tabel 6. Hasil Normalisasi

Sebelum	Sesudah
iklan nya banyak bgt sangat tidak nyaman, dikit dikit harus premium	iklan nya banyak banget sangat tidak nyaman dikit dikit harus premium
ngapa dah, mau muter musik aj harus premium, gk jelas, padahal kemarin gk	ngapa dah mau muter musik aja harus premium tidak jelas padahal kemarin tidak

e. *Stopword Removal*

Selanjutnya tahapan *stopword*, tahapan ini ialah menghapus kata stop, ini dimaksudkan untuk menghapus kata-kata umum yang tidak memiliki makna penting dan sering kali muncul, contoh dari kata stop antara lain “ini”, “nya” dan “tetap”. Hasil *stopword* tersebut dapat dilihat pada tabel 7.

Tabel 7. Hasil *Stopword Removal*

Sebelum	Sesudah
makin kikir, shuffle nya ga ngotak coba bisa <i>download</i> untuk di putar <i>online</i>	makin kikir shuffle nya ngotak coba <i>download</i> putar <i>online</i>

f. *Stemming*

Pada proses stemming, semua kata dengan imbuhan akan diubah menjadi kata dasar dalam proses ini. Hasil *stemming* dapat dilihat pada tabel 8.

Tabel 8. Hasil *Stemming*

Sebelum	Sesudah
playlist kadang suka ga bisa diputar dan ilang, gajelas	playlist kadang suka putar ilang gajelas
bagus tapi kurang di saat lagu yang di play malah lagu lain yang muncul	bagus kurang lagu play malah lagu muncul

3.3 TF-IDF

Setelah melewati *Preprocessing*, selanjutnya adalah tahapan "TF-IDF", sistem ini dirancang untuk menghitung bobot kata berdasarkan frekuensi kemunculannya dalam dokumen dan kelangkaannya di seluruh koleksi dokumen. Ini memungkinkan peneliti melihat kata-kata yang paling relevan atau memiliki bobot tertinggi dalam masing-masing dokumen, membantu dalam menganalisis fitur teks untuk tahap selanjutnya. Dalam tahap ini dapat dilihat pada tabel 9.

Tabel 9. Hasil Perhitungan TF

Kata	Skor TF
lengkap	1.0000
ngotak	0.2000
makin	0.2000
kikir	0.2000
shuffle	0.2000
nya	0.2000
mantap	1.0000

lirik	0.2500
sangat	0.2500
nya	0.2500
bagus	0.2500

Tabel 10 Hasil Perhitungan IDF

Kata	Skor IDF
lengkap	3.8691
ngotak	1.2343
makin	0.8452
kikir	0.7738
shuffle	0.7548
nya	0.4257
mantap	3.0807
lirik	0.8832
sangat	0.6402
nya	0.5322
bagus	0.3633

Tabel 11. Hasil Perhitungan TF-IDF

Kata	Skor TF-IDF
lengkap	3.8691
ngotak	1.2343
makin	0.8452
kikir	0.7738
shuffle	0.7548
nya	0.4257
mantap	3.0807
lirik	0.8832
sangat	0.6402
nya	0.5322
bagus	0.3633

3.4 Pengembangan Model

Pada tahapan pengembangan model sistem ini berfokus pada pembangunan dan pengujian model klasifikasi Naïve Bayes. Model *Naïve Bayes* dilatih dan diuji menggunakan perbandingan data 70% untuk pelatihan dan 30% untuk pengujian. Hasil pengembangan model akan ditampilkan mencakup Akurasi Model (77.24%) dalam bentuk persentase, selain itu ditampilkan juga diagram distribusi data dan tabel rincian hasil klasifikasi yang membandingkan Label Aktual dengan Label Prediksi pada data uji. Dalam tahap ini dapat dilihat pada tabel 12.

Tabel 12. Pengembangan Model

Teks Asli (Data Uji)	Label Aktual	Label Prediksi
Koleksi lagunya lumayan lengkap, jadi saya jarang tidak menemukan apa yang dicari	NETRAL	NETRAL
Harga langganan keluarga lumayan hemat dan terjangkau.	NETRAL	NETRAL

Teks Asli (Data Uji)	Label Aktual	Label Prediksi
Sejauh ini tidak ada yang salah dengan Spotify, ia hanya berfungsi sebagaimana mestinya.	NETRAL	NETRAL
bagus banget aku suka semua lagu ada semua penyanyi ada pokoknya i love spotify	POSITIF	POSITIF
bagusss	POSITIF	POSITIF

3.5 Teknik Pengujian

Tahap teknik pengujian ini merinci hasil evaluasi performa model. Menampilkan persentase hasil prediksi yang menunjukkan distribusi tebakan model pada data uji. Tabel *Confusion Matrix* memberikan analisis mendalam dengan membandingkan prediksi dan label aktual, sehingga terlihat jelas jumlah data yang benar terklasifikasi dan yang salah. Ditampilkan juga laporan klasifikasi yang merinci metrik seperti *precision*, *recall*, dan *f1-score* untuk setiap kelas sentimen. Dalam tahap ini dapat dilihat pada tabel 13.

Tabel 13. Teknik Pengujian

Label Prediksi	Persentase	Jumlah Data
NEGATIVE	26.90%	39
NETRAL	13.10%	19
POSITIF	60.00%	87

3.6 Pengujian

Bagian ini menjelaskan proses pengujian model klasifikasi sentimen yang telah dilatih menggunakan algoritma *Naïve Bayes*. Pengujian dilakukan untuk mengevaluasi performa dan efektivitas model dalam memprediksi sentimen ulasan pengguna Spotify. Dengan memilih rasio perbandingan 70% data pelatihan dan 30% data pengujian, lalu menekan tombol "Latih Model", sistem kemudian akan melatih model *Naïve Bayes* menggunakan data latih yang telah ditentukan. Setelah model terlatih, hasil pengembangan model akan ditampilkan mencakup Akurasi Model (73.79%) dalam bentuk persentase.

Hasil pengujian ini dirancang untuk memberikan evaluasi model yang lebih lengkap, dimulai dari ringkasan Persentase Hasil Prediksi yang memberikan gambaran umum tentang bagaimana model cenderung mengklasifikasikan data.. Kemudian, menampilkan tabel *Detail Confusion Matrix* yang memungkinkan analisis lebih mendalam, karena tabel ini memetakan secara rinci di mana saja model berhasil membuat prediksi yang tepat. Hasil *Detail Confusion Matrix* dapat dilihat pada tabel 14.

Tabel 14. Hasil *Detail Confusion Matrix*

AKTUAL/PREDIKSI	NEGATIVE	NETRAL	POSITIF
NEGATIVE	29	1	13
NETRAL	8	17	13
POSITIF	2	1	61

TP+TN+FP+FN untuk semua kelas) = 29+1+13+8+17+13+2+1+61 = 145 data.

Jumlah prediksi yang benar = 29(Negative)+17(Neutral)+61(Positif) = 107 data.

$$Akurasi = \frac{Jumlah\ Prediksi\ Benar}{Total\ Data\ Sampel\ Uji}$$

$$Akurasi = \frac{107}{145} = 73.79\%$$

Performa model kemudian diukur lebih lanjut secara kuantitatif melalui Laporan Klasifikasi, sebuah tabel standar yang menyajikan metrik penting seperti *Precision*, *Recall*, dan *F1-Score* untuk setiap kelas. Hasil Laporan Klasifikasi dapat dilihat pada tabel 15.

Tabel 15. Hasil Laporan Klasifikasi

UKURAN	PRECISION	RECALL	F1-SCORE	SUPPORT
NEGATIVE	0.74	0.67	0.71	43
NETRAL	0.89	0.45	0.60	38
POSITIF	0.70	0.95	0.81	64

Negative :

True Positif (TP_Negative) = Data aktual Negative yang di prediksi Negative = 29

False Negatif (FN_Negative) = Data aktual Negative yang di prediksi bukan Negative (yaitu Netral dan Positif) = 1+13 = 14

False Positif (FN_Negative) = Data aktual bukan Negative (yaitu Netral atau Positif) yang di prediksi Negative = 8+2 = 10

Support = TPnegative + FNnegative = 29+14 = 43

Precision (Negative) = $\frac{29}{29+10} = \frac{29}{39} = 0.74$

Recall (Negative) = $\frac{29}{29+14} = \frac{29}{43} = 0.67$

F1-Score (Negative) = $\frac{2 \times 0.7435 \times 0.6744}{0.7435 + 0.6744} = \frac{1.0024}{1.4179} = 0.71$

Netral :

True Positif (TP_Netral) = Data aktual Netral yang di prediksi Netral = 17

False Negatif (FN_Netral) = Data aktual Netral yang di prediksi bukan Netral (yaitu Negatif dan Positif) = 8+13 = 21

False Positif (FN_Netral) = Data aktual bukan Negative (yaitu Netral atau Positif) yang di prediksi Negative = 1+1 = 2

Support = TPnetral + FNnetral = 17+21 = 38

Precision (Negative) = $\frac{17}{17+2} = \frac{17}{19} = 0.89$

Recall (Negative) = $\frac{17}{17+21} = \frac{17}{38} = 0.45$

F1-Score (Negative) = $\frac{2 \times 0.8947 \times 0.4474}{0.8947 + 0.4474} = \frac{0.8010}{1.3421} = 0.560$

Positif :

True Positif (TP_Positif) = Data aktual Positif yang di prediksi Positif = 61

False Negatif (FN_Positif) = Data aktual Netral yang di prediksi bukan Netral (yaitu Negatif dan Positif) = 2+1 = 3

False Positif (FN_Positif) = Data aktual bukan Negative (yaitu Netral atau Positif) yang di prediksi Negative = 13+13 = 26

Support = TPnetral + FNnetral = 61+3 = 64

Precision (Negative) = $\frac{61}{61+26} = \frac{61}{86} = 0.70$

Recall (Negative) = $\frac{61}{61+3} = \frac{61}{64} = 0.95$

F1-Score (Negative) = $\frac{2 \times 0.7011 \times 0.9531}{0.7011 + 0.9531} = \frac{1.3364}{1.6542} = 0.81$

Weighted Average Precision :

WA_Precision = $(0.7435 \times \frac{43}{145}) + (0.8947 \times \frac{38}{145}) + (0.7011 \times \frac{64}{145})$

WA_Precision = $(0.7435 \times 0.29655) + (0.8947 \times 0.26207) + (0.7011 \times 0.44138)$

WA_Precision = $(0.2205) + (0.2348) + (0.3094)$

WA_Precision = 0.7647

Weighted Average Recall :

WA_Recall = $(0.6744 \times \frac{43}{145}) + (0.4474 \times \frac{38}{145}) + (0.9531 \times \frac{64}{145})$

WA_Recall = $(0.6744 \times 0.29655) + (0.4474 \times 0.26207) + (0.9531 \times 0.44138)$

WA_Recall = $(0.2000) + (0.1174) + (0.4207)$

WA_Recall = 0.7381

Weighted Average F1-Score :

$$WA_F1-Score = (0.7070 \times \frac{43}{145}) + (0.5968 \times \frac{38}{145}) + (0.8079 \times \frac{64}{145})$$

$$WA_F1-Score = (0.7070 \times 0.29655) + (0.5968 \times 0.26207) + (0.8079 \times 0.44138)$$

$$WA_F1-Score = (0.2098) + (0.1564) + (0.3567)$$

$$WA_F1-Score = 0.7229$$

4. KESIMPULAN

Berdasarkan hasil penelitian dan pembahasan analisis sentimen terhadap ulasan pengguna Spotify menggunakan algoritma Naïve Bayes, dapat disimpulkan beberapa poin penting sebagai berikut. Penelitian ini berhasil menerapkan algoritma *Naïve Bayes* untuk mengklasifikasikan ulasan pengguna Spotify ke dalam kategori sentimen positif, negatif, atau netral. Melalui 6 tahapan *preprocessing* seperti *case folding*, *cleansing*, normalisasi, *tokenization*, *stopword removal*, dan *stemming* terbukti efektif dalam menyiapkan data teks. Penerapan metode TF-IDF membantu dalam memberikan bobot penting pada kata-kata dalam ulasan, sehingga model dapat lebih fokus pada kata-kata kunci yang relevan dan memiliki makna penting dalam menentukan sentimen. Model *Naïve Bayes* menunjukkan kemampuan yang baik dalam menganalisis sentimen ulasan, dengan akurasi yang diperoleh mencapai 73.79%. Hal ini didukung oleh evaluasi menggunakan *confusion matrix* dan laporan klasifikasi yang menunjukkan performa model dalam memprediksi ulasan. Sistem yang dikembangkan memungkinkan pengguna untuk mengunggah data sendiri, melakukan proses *preprocessing* langkah demi langkah, menghitung TF-IDF, serta melatih model dengan rasio data latih dan uji yang dapat diatur secara dinamis. Ini membuat proses analisis sentimen menjadi lebih transparan dan mudah dipahami. Hasil analisis sentimen ini memberikan gambaran tentang persepsi pengguna terhadap Spotify sehingga dapat meningkatkan kualitas layanan dan pengalaman pengguna di masa mendatang.

DAFTAR PUSTAKA

- [1] D. Noviani, R. Pratiwi, S. Silvianadewi, M. Benny Alexandri, And M. Aulia Hakim, "Pengaruh Streaming Musik Terhadap Industri Musik Di Indonesia," *J. Bisnis Strateg.*, Vol. 29, No. 1, Pp. 14–25, 2020, Doi: 10.14710/Jbs.29.1.14-25.
- [2] A. S. Rahayu, A. Fauzi, And R. Rahmat, "Komparasi Algoritma Naïve Bayes Dan Support Vector Machine (Svm) Pada Analisis Sentimen Spotify," *J. Sist. Komput. Dan Inform.*, Vol. 4, No. 2, P. 349, 2022, Doi: 10.30865/Json.V4i2.5398.
- [3] C. B. Habibi And I. Irwansyah, "Konsumsi Dan Produksi Musik Digital Pada Era Industri Kreatif," *Metacommunication J. Commun. Stud.*, Vol. 5, No. 1, P. 23, 2020, Doi: 10.20527/Mc.V5i1.7449.
- [4] Apjii, "Pengguna Internet Di Daerah Tertinggal Tahun 2024," 2024.
- [5] W. P. Lesmana And A. Wijaya, "Analisis Sentimen Menggunakan Metode Naive Bayes Dan Support Vector Machine Pada Ulasan Aplikasi Joox Music," *J. Ilm. Sain Dan Teknol.*, Vol. 1, No. 2, Pp. 110–120, 2023.
- [6] G. Ginabila And A. Fauzi, "Analisis Sentimen Terhadap Pemutar Musik Online Spotify Dengan Algoritma Naive Bayes Dan Support Vector Machine," *J. Ilm. Ilk. - Ilmu Komput. Inform.*, Vol. 6, No. 2, Pp. 111–122, 2023, Doi: 10.47324/Ilkominfo.V6i2.180.
- [7] M. Azhary, "Analisis Sentimen Terhadap Pengguna Aplikasi Shopee Pada Twitter Menggunakan Algoritma Naïve Bayes," Pp. 1–23, 2025.
- [8] D. R. Firmansyah And E. Lestariningsih, "Analisis Sentimen Ulasan Aplikasi Smart Campus Unisbank Di Google Playstore Menggunakan Algoritma Naive Bayes," *J. Jtik (Jurnal Teknol. Inf. Dan Komunikasi)*, Vol. 8, No. 2, Pp. 498–507, 2024, Doi: 10.35870/Jtik.V8i2.1882.
- [9] R. Adhiyasya, S. I. Simanjuntak, A. B. Bali, And M. F. Fadholi, "Evaluasi Sentimen Pengguna Terhadap Aplikasi Bank Saqu Dengan Metode Algoritma Naïve Bayes," Vol. 7, No. 1, Pp. 1–12, 2025.
- [10] K. S. Putri, I. R. Setiawan, And A. Pambudi, "Analisis Sentimen Terhadap Brand Skincare Lokal Menggunakan Naïve Bayes Classifier," *Technol. J. Ilm.*, Vol. 14, No. 3, P. 227, 2023, Doi: 10.31602/Tji.V14i3.11259.
- [11] Syahril Dwi Prasetyo, Shofa Shofiah Hilabi, And Fitri Nurapriani, "Analisis Sentimen Relokasi Ibukota Nusantara Menggunakan Algoritma Naïve Bayes Dan Knn," *J. Komtekinfo*, Vol. 10, Pp. 1–7, 2023, Doi: 10.35134/Komtekinfo.V10i1.330.
- [12] M. Artama, I. N. Sukajaya, And G. Indrawan, "Classification Of Official Letters Using Tf-Idf Method," *J. Phys. Conf. Ser.*, Vol. 1516, No. 1, 2020, Doi: 10.1088/1742-6596/1516/1/012001.
- [13] D. E. Cahyani And I. Patasik, "Performance Comparison Of Tf-Idf And Word2vec Models For Emotion Text Classification," *Bull. Electr. Eng. Informatics*, Vol. 10, No. 5, Pp. 2780–2788, 2021, Doi: 10.11591/Eei.V10i5.3157.