

## ANALISIS SENTIMEN PENGGUNA APLIKASI SHEJEK BERDASARKAN ULASAN DI GOOGLE PLAY STORE MENGGUNAKAN METODE NAÏVE BAYES

Fatmawati, Bambang Irawan, Agus Bahtiar

Program Studi Teknik Informatika, STMIK IKMI Cirebon

Jl. Perjuangan No.10 B, Kota Cirebon

*pfatmawati420@gmail.com*

### ABSTRAK

Di era digital dewasa ini, perkembangan aplikasi transportasi online telah menjadi bagian tak terpisahkan dari kehidupan sehari-hari pada masyarakat, menyebabkan perubahan relevan dalam melakukan perjalanan, *Shejek* menawarkan layanan transportasi daring bagi pengguna untuk memesan kendaraan perjalanan mereka. popularitas *Shejek* tidak selalu diikuti ulasan yang positif dari pengguna. Permasalahan muncul dalam mencari cara secara efektif memperoleh opini pengguna dari ulasan yang ditinggalkan di *Google Play Store* terkait dengan aplikasi *Shejek*. Tujuannya adalah untuk mengembangkan model analisis sentimen yang dapat membantu memahami masukan pengguna secara lebih efektif, Metode yang digunakan dalam penelitian ini adalah analisis sentimen menggunakan algoritma *Naive Bayes*. Penelitian ini bertujuan untuk menganalisis tingkat akurasi, precision, dan recall dalam penggunaan algoritma *Naive Bayes* mengklasifikasikan sentimen ulasan aplikasi *Shejek* di *Google Play Store*. Data yang digunakan adalah ulasan pengguna yang diambil dari *Google Play Store* terkait aplikasi *Shejek* pada 627 ulasan setelah *preprocessing* data. Hasilnya menunjukkan 500 ulasan dengan label, terdiri dari 271 ulasan positif dan 229 ulasan negatif. Analisis distribusi sentimen menggambarkan mayoritas ulasan cenderung bersifat positif. Evaluasi performa model menunjukkan tingkat akurasi sebesar 80.39%, precision sebesar 73.69%, dan recall sebesar 94.54%. Hasil ini memberikan wawasan terkait aplikasi *Shejek* oleh pengguna, sekaligus memberikan pengembangan dan perbaikan lebih lanjut.

**Kata Kunci :** Analisis sentiment, Aplikasi Shejek, Algoritma Naive Bayes, Transportasi Online

### 1. PENDAHULUAN

Di era teknologi informasi saat ini, aplikasi mobile menjadi kebutuhan penting dalam keseharian. *Shejek* sebagai inovasi di bidang transportasi online menyediakan layanan untuk memudahkan aktivitas pengguna *Google Play Store* sebagai platform utama distribusi aplikasi android menjadi wadah bagi pengguna menyampaikan pandangan melalui ulasan. Memahami sentiment dan opini pengguna tentang aplikasi *Shejek* di platform ini penting untuk pengembangan dan peningkatan kualitas layanannya analisis sentiment merupakan pendekatan efektif untuk mengeksplorasi dan menginterpretasikan tanggapan pengguna terhadap *Shejek*, dengan menerapkan algoritma *Naive Bayes* hasil klasifikasi sentiment bertujuan membantu *Shejek* dalam mengevaluasi persepsi masyarakat umum.

Pemahaman sentiment pengguna terhadap aplikasi *Shejek* menjadi dasar utama penelitian ini Analisis sentiment dipilih sebagai metode untuk menggali persepsi, kepuasan, atau kritik yang terkandung dalam ulasan pengguna *Shejek* dalam konteks kajian ini algoritma *Naive Bayes* diimplementasikan untuk mengklasifikasikan pemahaman mendalam terkait preferensi dan pengalaman pengguna aplikasi *Shejek*. Penelitian ini tidak hanya bermanfaat bagi pengembang *Shejek* tetapi juga memberi wawasan yang berharga bagi peminat kepentingan lainnya seperti pihak pemasaran, layanan pelanggan, dan pengambilan keputusan strategis. Dengan memahami sentiment masyarakat terhadap aplikasi ini diharapkan dapat

ditemukan pola umum, kekuatan, dan kelemahan yang dapat membimbing perbaikan dan pengembangan lebih lanjut.

Tujuan dari penelitian untuk menganalisis sentiment (pendapat /opini) dari ulasan pengguna aplikasi *Shejek* di *Google Play Store* dengan menggunakan algoritma *Naive Bayes* penelitian ini bertujuan mengidentifikasi dan memahami sentiment umum pada ulasan pengguna *Shejek*, guna memberikan wawasan tentang bagaimana masyarakat merespons dan memperspektifkan aplikasi ini secara keseluruhan, penelitian akan menganalisis pola sentiment positif dan negative yang muncul pada ulasan komentar untuk mengekstrak informasi terkait kekuatan dan kelemahan *Shejek* dari perspektif pengguna tujuan lain adalah mengidentifikasi masalah *Shejek* yang paling berpengaruh terhadap sentiment agar pengembang dapat berfokus pada peningkatan yang memerlukan perhatian lebih. Secara keseluruhan penelitian ini dimaksudkan untuk memberi wawasan mendalam bagi *Shejek* agar dapat membuat keputusan yang lebih baik terkait inovasi fitur perbaikan layanan dan strategi pemasaran yang lebih efektif berdasarkan pemahaman atas sentiment dan harapan pengguna serta masyarakat umum, dengan demikian diharapkan dapat meningkatkan kualitas layanan dan kepuasan melalui inovasi yang lebih terhadap kebutuhan masyarakat selain itu, penelitian ini dapat membuka wawasan bagi pengembang lebih lanjut dalam analisis sentiment ulasan terhadap opini masyarakat terkait pengguna

aplikasi Shejek di *Google Play Store* menggunakan algoritma Naïve Bayes.

Metode pendekatan yang digunakan dalam penelitian ini adalah kuantitatif dengan jenis penelitian deskriptif untuk menganalisis data ulasan pengguna Shejek, teknik text mining dan algoritma Naïve Bayes Classifier dipilih sebagai metode analisis sentiment yang akan mengklasifikasi data teks ulasan ke dalam dua kelas sentiment yaitu positif dan negative. Hasil klasifikasi tersebut diharapkan dapat mendeskripsikan distribusi dan tren sentiment public pengguna Shejek berdasarkan ulasan yang ditinggal di *Google Play Store*.

## 2. TINJAUAN PUSTAKA

Dari penelitian sebelumnya yang dilakukan Isan, dkk. (2023) membahas mengenai hasil penelitian analisis sentiment pengguna aplikasi Brimo kedalam kategori positif dan negative dimana data dikumpulkan melalui hasil pengumpulan data pada bulan Agustus 2022 hingga Januari 2023 dengan jumlah data teks sebanyak 1.550, hasilnya menunjukkan bahwa pengguna aplikasi memberikan ulasan positif sehingga menghasilkan prediksi jumlah hasil klasifikasi sentiment terdapat 1.012 komentar positif dan 894 komentar negative pada data teks yang berarti banyak pengguna yang memberikan tanggapan positif, umpan balik negative untuk aplikasi Brimo dan hasil klasifikasi menggunakan metode Naïve Bayes menunjukkan akurasi sebesar 84,52%, presisi sebesar 82,51%, dan recall sebesar 87,62% [1].

Riset oleh Ramadhan, dkk. (2023) membahas mengenai perkembangan E-commerce di Indonesia telah memunculkan platform marketplace seperti Tokopedia, Shopee, dan Lazada yang sudah diunduh lebih dari 100 juta pengguna, akan tetapi banyak opini dan tanggapan yang kontroversial dari masyarakat terhadap ketiga aplikasi tersebut maka dari itu penelitian ini akan melakukan analisis terhadap data ulasan kritik dari pengguna. Banyak pendapat yang mendukung dan menentang aplikasi ini peneliti akan melakukan analisis sentiment terhadap data ulasan pengguna ketiga marketplace untuk memahami sentiment positif atau dominan negative, dengan demikian hasil sentiment dapat menjadi informasi berharga bagi pengguna dalam memilih rekomendasi marketplace online terbaik di di Indpnesia berdasarkan opini dan ulasan Masyarakat [2].

Husnina, dkk. (2023) menggunakan metode Naïve Bayes untuk mengklasifikasikan ulasan aplikasi RedBus dengan memanfaatkan dataset yang diperoleh dari *Google Play Store* sebanyak 500 ulasan. Hasil penelitian menunjukkan bahwa penggunaan algoritma Naïve Bayes dengan pembobotan TF-IDF dan pengujian cross-validation dengan nilai k=10 menghasilkan akurasi terbaik sebesar 93,56% [3].

### 2.1. Data Mining

Penambangan data (*Data Mining*) adalah proses menganalisis dan mengekstraksi pengetahuan secara

otomatis menggunakan satu atau lebih algoritma pembelajaran mesin. Demi mendukung proses pengambilan keputusan di masa yang akan datang, penambangan data melibatkan identifikasi tren atau pola yang diinginkan dalam basis data yang berskala besar. Alat-alat tertentu mengidentifikasi pola-pola ini, dan alat-alat tersebut dapat memberikan analisis data yang bermanfaat dan perseptif yang dapat diperiksa lebih lanjut, mungkin dengan bantuan alat pendukung keputusan tambahan [4].

### 2.2. Analisis Sentimen

Analisis Sentimen adalah evaluasi dari aspek emosional dalam teks subjektif yang melibatkan analisis, pemrosesan, rangkuman, dan inferensi. Saat ini, Analisis sentiment terbagi menjadi dua pendekatan utama, yaitu penggunaan klasifikasi pembelajaran mesin dan metode klasifikasi berdasarkan aturan. Pendekatan pembelajaran mesin menggunakan kata-kata emosi sebagai fitur klasifikasi, sementara kamus emosi dapat digunakan untuk mengidentifikasi ciri-ciri sentimen dengan cepat dan efisien. [5].

### 2.3. Confusion Matrix

*Confusion Matrix* merupakan suatu metode yang digunakan untuk mengevaluasi performa suatu sistem dengan membandingkan hasil prediksi sistem dengan data aktual yang ada [6]. Presentasi dari *Confusion Matrix* dapat dilihat melalui Tabel 1.

Tabel 1. Confusion matrix

<i>Predicted Value</i>	<i>Actual Value</i>	
	<i>Positive</i>	<i>Negative</i>
<i>Positive</i>	TP	FP
<i>Negative</i>	FN	TN

Keterangan Tabel:

TP : Data positif yang diprediksi positif

TF : Data negatif yang diprediksi negatif

FP : Data negatif yang diprediksi positif

FN : Data positif yang diprediksi negatif

*Confusion Matrix* menghasilkan keluaran berupa *F1-score*, yang merupakan nilai rata-rata harmonik dari presisi dan *recall*. Akurasi digunakan untuk mengukur sejauh mana metode tersebut mampu memprediksi output dengan benar. *Recall* mencerminkan rasio antara jumlah sampel positif yang terdeteksi dengan benar terhadap jumlah total sampel yang sebenarnya positif, sementara presisi adalah proporsi antara jumlah data positif yang berhasil diklasifikasikan secara akurat oleh sistem klasifikasi dengan jumlah keseluruhan data yang diklasifikasikan sebagai data positif oleh sistem klasifikasi. Setiap keluaran ini dinyatakan dalam bentuk persentase, dengan rentang nilai antara 1 hingga 100%. Berikut adalah rumus yang digunakan untuk menghitung nilai dari masing-masing keluaran *Confusion Matrix* [7].

### 2.4. Naïve Bayes

Algoritma Naïve Bayes merupakan suatu bentuk klasifikasi yang memperhitungkan nilai probabilitas statistik. Oleh karena itu, prediksi keanggotaan suatu kelas data baru dilakukan dengan memprediksi probabilitas keanggotaan data tersebut [8]. Persamaan teorema Bayes diberikan pada persamaan

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)} \tag{1}$$

Keterangan:

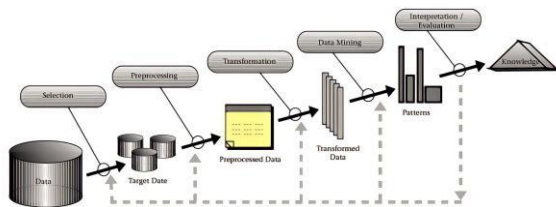
A = Kelas prediksi (positif, negatif, netral)

B = Data ulasan

P(A|B) = Probabilitas kelas A berdasarkan data B

### 3. METODE PENELITIAN

Penelitian ini menggunakan metode analisa Knowledge Discovery in Database (KDD). Tahapan analisa data dapat dilihat pada gambar berikut :



Gambar 1. Tahapan KDD

Knowledge Discovery in Databases (KDD) merujuk pada serangkaian proses yang digunakan untuk menemukan pengetahuan yang dapat memberikan manfaat dari suatu kumpulan data yang melibatkan rangkaian langkah transformasi, termasuk tahapan preprocessing dan post-processing [9].

#### 3.1. Data selection

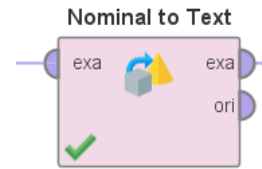
Data selection adalah proses pemilihan subset data dari kumpulan data yang lebih besar berdasarkan kriteria atau parameter tertentu. Ini sering dilakukan untuk fokus pada aspek data tertentu atau untuk mengurangi jumlah data yang perlu dianalisis atau diproses. Data diambil dari website google play store langsung dengan menggunakan teknik scraping dan diperoleh data mentah sebanyak 500 ulasan.

#### 3.2. Data Preprocessing

Preprocessing dilakukan untuk mengidentifikasi bagian data yang tidak lengkap atau tidak akurat ataupun data yang hilang.

#### 3.3. Transformation

Transformasi dilakukan untuk mengubah atribut yang memiliki nilai nominal kemudian diganti menjadi text.



Gambar 2. Operator Nominal to Text

Operator ini biasanya diterapkan pada variabel target yang akan dianalisis sentimennya.

### 3.4. Data Mining

Tujuan dari data mining adalah untuk menerapkan teknik dan metode tertentu untuk mengekstrak informasi yang berarti dari sejumlah besar data. Dalam penelitian ini, menerapkan algoritma Naïve Bayes untuk mengukur dan mengevaluasi kinerja algoritma. Proses selanjutnya adalah klasifikasi algoritma Naïve Bayes.

### 3.5. Evaluation

Evaluation didefinisikan untuk mengidentifikasi model yang meningkatkan wawasan berdasarkan metrik tertentu.

## 4. HASIL DAN PEMBAHASAN

Peneliti akan menyajikan hasil dan analisis yang diperoleh selama tahap penelitian berlangsung. Hasil penelitian ini akan menjelaskan secara terperinci Langkah-langkah yang dilakukan pada saat pengumpulan data, preprocessing data, transformation data, dan evaluasi data.

#### 4.1. Selection

Proses ini melakukan pembacaan data dalam bentuk file csv dengan menggunakan operator Read CSV seperti pada gambar 2 berikut.



Gambar 3. Operator Read CSV

Dari hasil pembacaan operator Read CSV didapatkan informasi data seperti pada gambar 4 dibawah ini.

Ulasan	Label
polynomial	binomial
1 Saya senang ada apk khusus untuk perempuan semoga kedepannya bisa lebih banyak	Positif
2 Assalamu'alaikum warahmatullahi wabarakatuh. Mohon maaf mau tanya apakah apka	Positif
3 Apk nya bagus bisa pesan lewat admin. Semoga lebih banyak lagi diversinya sangat me	Positif
4 Assalamu'alaikum, min saya buka aplikasinya ko sudah sih... padahal udah daftar... to ga bis	Positif
5 Ini lokasiya sudah tersebar dimana saja, ya? Terus mau nyoba pesen kok gak bisa ml	Positif
6 Untuk wilayah surabaya apa sudah ada? Bagaimana car mendaftar jadi sheikh driver ul	Positif
7 Aplikasinya bagus khusus untuk perempuan jadi lebih nyaman dengan sesama peremp	Positif
8 Alhamdulillah ada aplikasi Shehik, aplikasi ini sangat membantu bagi muslimah yang ber	Positif
9 Semoga Shehik tambah banyak dan diterima di masyarakat, yang khususnya perempuan	Positif
10 Keren banget ni aplikasi, jadi solusi wanita yg nggamau dibonceng sama laki laki. Su	Positif
11 Keren... Shehik apakah bisa langganan bukannya...? oh anak perempuan juga kuliah di pgg	Positif
12 Berhenti ko diproteksi system di service lbih gampang sah ke lebih bagus dan meria	Positif
13 Iya laki saya tau apk nya tidak dapat di pakai... dia ngejain setiap pesen 10k ke admin ta	Positif
14 masya Allah sy senang nalnya klu mau pergi jauh sy g betera biasa motor sendiri.5	Positif
15 Alwin ini sudah ada di daerah jakarta kah? Mau coba pesen kok tidak bisa pilih lokasi y	Positif
16 Subhanallah membantu para perempuan yg ingin pergi ke suatu tempat.5	Positif
17 Sangat nyaman km sesama perempuan tdk menimbulkan fitnah.5	Positif
18 Tolong hadiahkan juga untuk wilayah Balikpapan, Kalimantan Timur.4	Positif
19 Rausin Pk, trusmi bawabaha meningkatkan performa asa untuk membantu muslimah ta	Positif

Gambar 4. Data Ulasan

Dataset berikut didapatkan secara langsung dari *website play store* dengan melakukan teknik *scraping* menggunakan *google collab* dengan *API Google-Play-Scraper*. Langkah selanjutnya yaitu menambahkan operator *Set Role*. Operator *Set Role* di Rapidminer digunakan untuk mengubah peran dari satu atau lebih atribut. Peran atribut menentukan cara operator lain menangani atribut tersebut. Peran *default* adalah *regular*, peran lainnya diklasifikasikan sebagai *special*. Pada penelitian ini, Operator *Set Role* digunakan untuk memberikan label pada atribut sentimen yang bertindak sebagai atribut target operator learning.

4.2. Preprocessing

Pre-processing bertujuan untuk menghilangkan noise dan menyeragamkan bentuk kata untuk mengurangi volume kosakata. Langkah-langkah preprocessing seperti *Tokenize*, *Transform Cases*, *Filter Tokens (by Length)*, dan *Stopwords (Dictionary)*.

Tabel 2. Contoh tahap *transform case*

Sebelum	Sesudah
Tolong hadirkan juga untuk wilayah Balikpapan Kalimantan Timur	tolong hadirkan juga untuk wilayah balikpapan kalimantan timur

Pada Tabel 2. merupakan contoh hasil transformasi dari tahap *transform case*. Tahap ini untuk mengonversi semua huruf dalam teks ke huruf kecil atau besar agar tidak ada perbedaan dalam pemrosesan teks.

Tabel 3. Contoh tahap *tokenizing*

Sebelum	Sesudah
tolong hadirkan juga untuk wilayah balikpapan kalimantan timur	'tolong' 'hadirkan' 'juga' 'untuk' 'wilayah' 'balikpapan' 'kalimantan' 'timur'

Pada Tabel 3. merupakan contoh hasil transformasi dari tahap *Tokenizing*, Memecah teks menjadi token atau kata-kata individu agar dapat diolah lebih lanjut.

Tabel 4. Contoh tahap *token (by length)*

Sebelum	Sesudah
'tolong' 'hadirkan' 'juga' 'untuk' 'wilayah' 'balikpapan' 'kalimantan' 'timur'	'tolong' 'hadirkan' 'juga' 'untuk' 'wilayah' 'balikpapan' 'kalimantan' 'timur'

Pada Tabel 4. diberikan ilustrasi tentang implementasi tahap *Filter Token (By Length)*. Langkah ini bertujuan untuk mengeliminasi kata-kata yang dianggap tidak relevan setelah proses *Tokenizing*. Dalam pelaksanaannya, kata-kata dengan panjang tertentu akan dihapus. Dalam konteks ulasan, kata-kata yang memiliki panjang kurang dari 4 kata atau lebih dari 25 kata dihapus [10].

Tabel 5. Contoh tahap *stopwords*

Sebelum	Sesudah
'tolong' 'hadirkan' 'juga' 'untuk' 'wilayah' 'balikpapan' 'kalimantan' 'timur'	'tolong' 'hadirkan' 'juga' 'untuk' 'wilayah' 'balikpapan' 'kalimantan' 'timur'

Pada Tabel 5. merupakan contoh hasil transformasi dari tahap *Stopwords*. Pada tahap stopwords, kata-kata yang tidak penting akan dihilangkan untuk menghasilkan teks yang lebih bermakna dan *relevan* dengan analisis sentimen

4.2. Labeling

Data diberi label secara manual sebagai sentimen positif dan negatif.

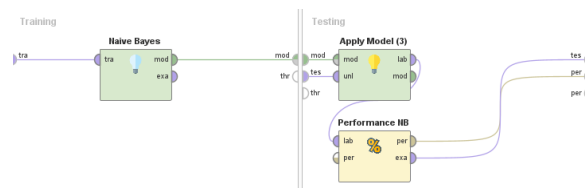
Tabel 6. Labeling

No.	Ulasan	Sentimen
1.	Alhamdulillah ada aplikasi Shejek, aplikasi ini sangat membantu bagi muslimah yang berpergian	Positif
2.	susah banget cari alamat nya sering ga ada mohon di tingkatkan lagi yaa	Negatif

Label adalah cara terbaik untuk memberikan informasi seperti ID, nama, lokasi, dan ulasan. Pada penelitian ini, sentimen ditentukan berdasarkan *rating*. *Rating* 5 berarti positif dan *rating* 1 berarti negatif.

4.2. Naïve Bayes

Proses selanjutnya adalah klasifikasi algoritma Naïve Bayes. Proses ini dilakukan untuk mengukur dan mengevaluasi kinerja algoritma. Selama fase ini, data diperiksa untuk menghindari pengelompokan data. Berikut adalah gambar dari proses implementasi algoritma Naïve Bayes.



Gambar 5. Desain evaluasi Naive Bayes

Naïve Bayes adalah algoritma yang terintegrasi dengan pembelajaran mesin yang dapat digunakan untuk klasifikasi atau regresi. Naïve Bayes adalah algoritma yang kuat dan serbaguna yang dapat diterapkan pada berbagai masalah, termasuk analisis sentimen. Operator *apply model* memiliki kemampuan untuk menerapkan model pembelajaran mesin ke data baru. Operator ini menggunakan model terlatih untuk memprediksi nilai atribut keluaran untuk data baru. Operator kinerja (*performance*) digunakan untuk mengevaluasi kinerja model pembelajaran mesin. Operator-operator ini menghasilkan berbagai metrik kinerja seperti akurasi, presisi, dan recall. Hasil performa yang diperoleh dengan menerapkan metode Naïve Bayes dapat dilihat pada gambar berikut.

	true Positif	true Negatif	class precision
pred Positif	273	30	90,10%
pred Negatif	130	373	90,10%
class recall	97,74%	90,90%	

Gambar 6. Performa Naïve Bayes

Berdasarkan Gambar 6 diatas, kinerja yang disiapkan pada subjek analisis sentimen terhadap review aplikasi Shejek di *Google Play Store* memiliki *accuracy* sebesar 80,17% *precision* sebesar 90,10% dan *recall* 92,56%.

4.2. Evaluation

Tujuan evaluasi adalah untuk mengevaluasi seberapa akurat model algoritma yang dibuat. Evaluasi kinerja ini dilakukan dengan mengutamakan akurasi sebagai parameter evaluasi untuk memperoleh nilai presisi dan recall, penelitian ini menggunakan *confusion matrix*. *Confusion matrix* berfungsi sebagai alat analisis untuk mengevaluasi kemampuan model klasifikasi dalam mengenali tupel data yang berbeda.

*True Positive* (TP) adalah kelas positif yang diprediksi secara tepat sebagai kelas positif. Pada penelitian ini diperoleh 273 kelas positif. *False Positif* (FP) merupakan kelas negatif yang diprediksi salah sebagai kelas positif diperoleh sebanyak 30. Kemudian *True Negative* (TN) merupakan kelas negatif yang diprediksi benar, diperoleh sebanyak 373 kelas negatif, dan *False Negative* (FN) merupakan kelas positif yang salah diprediksi diperoleh sebanyak 130 kelas negatif. Di bawah ini adalah hasil perhitungan yang dilakukan secara manual dengan menggunakan perhitungan *confusion matrix*.

Rumus perhitungan nilai *Accuracy*:

$$Accuracy = \frac{273 + 373}{273 + 30 + 130 + 373} = 80,14\%$$

Rumus perhitungan nilai *Precision*:

$$Precision = \frac{273}{273 + 30} = 90,10\%$$

Rumus perhitungan nilai *Recall*:

$$Recall = \frac{273}{273 + 130} = 67,74\%$$

Hasil penelitian ini menunjukkan keberhasilan metode Naïve Bayes dengan tingkat *accuracy* sebesar 80,14%. *Micro average* yang juga mencapai 80,15% memberikan gambaran keseluruhan tentang efisiensi model klasifikasi ini. Selain itu, analisis hasil *performance* mengungkapkan tingkat *precision* sebesar 90,10%, yang mencerminkan kemampuan model untuk mengidentifikasi dengan tepat ulasan yang benar-benar positif. *Recall* sebesar 67,74% menunjukkan sejauh mana model dapat mengenali seluruh ulasan yang sebenarnya positif. Dengan

demikian, temuan ini memberikan indikasi positif terhadap efektivitas Naïve Bayes dalam menganalisis sentimen ulasan aplikasi *Shejek*.

5. KESIMPULAN DAN SARAN

Pada penelitian ini, mengenai analisis sentimen pendekatan *machine learning* menggunakan algoritma Naïve Bayes pada ulasan pengguna aplikasi shejek, model yang dikembangkan dievaluasi dan diperoleh model terbaik dengan nilai akurasi sebesar 80,14%, tingkat presisi sebesar 90,10%, dan tingkat recall sebesar 67,74%. Peneliti dapat menyimpulkan bahwa algoritma Naïve Bayes adalah metode yang cocok untuk mencapai hasil yang baik dalam klasifikasi *data mining*. Penelitian di masa depan didorong untuk memperluas penggunaan metode klasifikasi dan optimasi yang berbeda untuk membandingkan kinerja dengan metode yang diterapkan sebelumnya. Selain itu, penelitian lebih lanjut dapat memperpanjang periode pengumpulan data dan membuat model yang dikembangkan menjadi lebih akurat.

DAFTAR PUSTAKA

- [1] M. Khoirul, U. Hayati, and O. Nurdiawan, "ANALISIS SENTIMEN APLIKASI BRIMO PADA ULASAN PENGGUNA DI GOOGLE PLAY MENGGUNAKAN ALGORITMA NAIVE BAYES," 2023.
- [2] B. Z. Ramadhan, I. Riza, and I. Maulana, "Analisis Sentimen Ulasan Pada Aplikasi E-Commerce Dengan Menggunakan Algoritma Naïve Bayes," 2022. [Online]. Available: <http://jurnal.polibatam.ac.id/index.php/JAIC>
- [3] D. Nazhifa, N. Husnina, D. E. Ratnawati, and B. Rahayudi, "Analisis Sentimen Pengguna Aplikasi RedBus berdasarkan Ulasan di Google Play Store menggunakan Metode Naïve Bayes," 2023. [Online]. Available: <http://j-ptiik.ub.ac.id>
- [4] Haris Kurniawan, Sarjon Defit, and Sumijan, "Data Mining Menggunakan Metode K-Means Clustering Untuk Menentukan Besaran Uang Kuliah Tunggal," *Journal of Applied Computer Science and Technology*, vol. 1, no. 2, pp. 80–89, Dec. 2020, doi: 10.52158/jacost.v1i2.102.
- [5] A. Deolika and E. Taufiq Luthfi, "ANALISIS PEMBOBOTAN KATA PADA KLASIFIKASI TEXT MINING," *Jurnal Teknologi Informasi*, vol. 3, no. 2, 2019.
- [6] A. Dwiki, A. Putra, and S. Juanita, "Analisis Sentimen Pada Ulasan Pengguna Aplikasi Bibit Dan Bareksa Dengan Algoritma KNN," vol. 8, no. 2, 2021, [Online]. Available: <http://jurnal.mdp.ac.id>
- [7] H. Hidayatullah and Y. Umaidah, "PENERAPAN NAÏVE BAYES DENGAN OPTIMASI INFORMATION GAIN DAN SMOTE UNTUK ANALISIS SENTIMEN PENGGUNA APLIKASI CHATGPT," 2023.
- [8] D. Normawati and S. A. Prayogi, "Implementasi Naïve Bayes Classifier Dan Confusion Matrix

- Pada Analisis Sentimen Berbasis Teks Pada Twitter,” 2021.
- [9] A. Rahman Isnain, A. Indra Sakti, D. Alita, and N. Satya Marga, “SENTIMEN ANALISIS PUBLIK TERHADAP KEBIJAKAN LOCKDOWN PEMERINTAH JAKARTA MENGGUNAKAN ALGORITMA SVM,” *JDMI*, vol. 2, no. 1, pp. 31–37, 2021, [Online]. Available: <https://t.co/NfhmfMjtXw>
- [10] L. Asri *et al.*, “ANALISIS SENTIMEN OPINI PUBLIK BERITA KEBAKARAN HUTAN MELALUI KOMPARASI ALGORITMA SUPPORT VECTOR MACHINE DAN K-NEAREST NEIGHBOR BERBASIS PARTICLE SWARM OPTIMIZATION,” *Jurnal Pilar Nusa Mandiri*, vol. 13, no. 1, 2017, [Online]. Available: [www.tribunnews.com](http://www.tribunnews.com)