Community Engagement & Emergence Journal
Volume 6 Nomor 6, Tahun 2025
Halaman : 4757-4776

Penerapan Metode Naive Bayes Untuk Klarifikasi Keluhan Pengguna Aplikasi MyTelkomsel

Application Of Naive Bayes Method For Clarifying Customer Complaints On MyTelkomsel Application

Glori Yani Hutabarat[1] Sintaria Sembiring[2]
Universitas Advent Indonesia Jl. Kolonel Masturi No.288, Bandung Barat[1,2]
Fakultas Teknologi Informasi, Universitas Advent Indonesia[1,2]
2281026@unai.edu[1] sintaria.sembiring@unai.edu[2]

Abstrak

Penelitian ini bertujuan untuk menganalisis keluhan pengguna terhadap aplikasi MyTelkomsel dengan menggunakan metode Naïve Bayes sebagai model klasifikasi teks. Permasalahan yang diangkat adalah banyaknya keluhan pengguna mengenai layanan jaringan, fitur aplikasi, dan harga paket yang tinggi. Penelitian ini menggunakan pendekatan kuantitatif dengan metode text mining dan machine learning berbasis algoritma Naïve Bayes. Subjek penelitian adalah 800 ulasan negatif pengguna yang dikumpulkan dari Google Play Store periode 1 Mei hingga 1 September 2025. Data diolah melalui tahapan cleaning, case folding, normalisasi, stopword removal, tokenisasi, stemming, dan pembobotan TF-IDF. Hasil pengujian menunjukkan akurasi 78%, precision 54,5%, recall 71,4%, dan F1-score 61,8%. Penelitian ini menyimpulkan bahwa Naïve Bayes efektif untuk mengidentifikasi kategori keluhan pengguna. Implikasi hasil penelitian ini diharapkan dapat membantu Telkomsel dalam meningkatkan kualitas layanan dan menjadi dasar pengembangan sistem analisis otomatis untuk pengelolaan umpan balik pelanggan.
Kata kunci: Naïve Bayes, Klasifikasi Keluhan, MyTelkomsel, TF-IDF, Analisis Teks

Abstract

This research aims to analyze user complaints about the MyTelkomsel application using the Naïve Bayes method as a text classification model. The problem raised was the large number of user complaints regarding network services, application features and high package prices. This research uses a quantitative approach with text mining and machine learning methods based on the Naïve Bayes algorithm. The research subjects were 800 negative user reviews collected from the Google Play Store for the period 1 May to 1 September 2025. The data was processed through the stages of cleaning, case folding, normalization, stopword removal, tokenization, stemming, and TF-IDF weighting. Test results show 78% accuracy, 54.5% precision, 71.4% recall, and 61.8% F1-score. This research concludes that Naïve Bayes is effective for identifying categories of user complaints. It is hoped that the implications of the results of this research will help Telkomsel improve service quality and become the basis for developing an automatic analysis system for managing customer feedback.
Keywords: *Naïve Bayes, Complaint Classification, MyTelkomsel, TF-IDF, Text Analysis*

1. Pendahuluan

Dalam era digital yang semakin maju, aplikasi seluler memegang peranan penting dalam memenuhi kebutuhan komunikasi masyarakat. Aplikasi self-service seperti MyTelkomsel memberikan kemudahan bagi pengguna dalam membeli paket data, mengecek pulsa, dan mengakses promo. Namun, meskipun MyTelkomsel merupakan salah satu aplikasi telekomunikasi terbesar di Indonesia, masih banyak pengguna yang mengeluhkan performa aplikasi, harga paket, dan kualitas jaringan. Hal ini terlihat dari banyaknya ulasan negatif pada Google Play Store, yang

e-ISSN:2715-9752, p-ISSN:2715-9868
Copyright © 2025 THE AUTHOR(S). This article is distributed under a Creative Commons Attribution-NonCommercial 4.0 International license

mencerminkan adanya kesenjangan antara harapan dan pengalaman pengguna (service quality gap).

Beberapa penelitian sebelumnya telah membahas analisis sentimen pada aplikasi e-commerce dan layanan digital menggunakan Naïve Bayes [4][5], namun belum banyak yang secara spesifik mengkaji klasifikasi keluhan pada aplikasi telekomunikasi berbasis data ulasan aktual. Inilah gap yang ingin dijawab oleh penelitian ini. Dengan memahami pola keluhan pengguna secara sistematis, Telkomsel dapat meningkatkan layanan yang lebih responsif dan sesuai kebutuhan pelanggan.

Penelitian ini memiliki kebaruan dalam penerapan Naïve Bayes untuk mengidentifikasi kategori keluhan pengguna MyTelkomsel berdasarkan data aktual dari Google Play Store. Tujuan utama penelitian ini adalah mengklasifikasikan jenis keluhan pengguna dan mengevaluasi kinerja model menggunakan metrik akurasi, precision, recall, dan F1-score. Selain itu, penelitian ini diharapkan memberikan kontribusi pada pengembangan sistem analitik otomatis untuk mendukung peningkatan pengalaman pengguna.

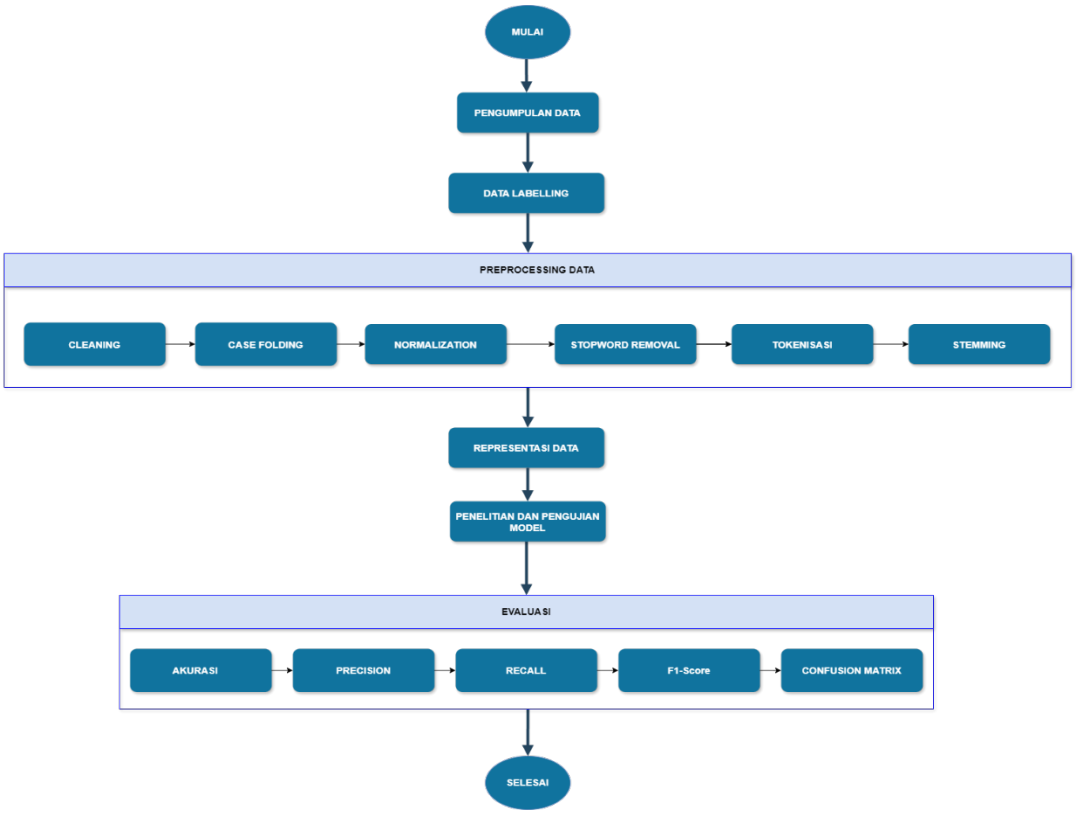
2. Metode

Penelitian ini menggunakan pendekatan kuantitatif dengan metode text classification berbasis Naïve Bayes. Subjek penelitian adalah 800 ulasan negatif pengguna aplikasi MyTelkomsel yang diambil melalui teknik web crawling menggunakan bahasa pemrograman Python di platform Google Colab. Data dikumpulkan antara 1 Mei hingga 1 September 2025.

Prosedur penelitian meliputi:

1. Pengumpulan Data: menggunakan Google Play Scraper untuk mengekstrak ulasan pengguna.
2. Preprocessing: meliputi cleaning, case folding, normalization, stopword removal, tokenization, dan stemming untuk memastikan data siap diproses.
3. Representasi Data: menggunakan metode TF-IDF untuk mengubah teks menjadi vektor numerik.
4. Klasifikasi: model Naïve Bayes digunakan untuk memprediksi kategori keluhan.
5. Evaluasi Model: dilakukan dengan confusion matrix, akurasi, precision, recall, dan F1-score.

Instrumen penelitian berupa skrip Python untuk pengumpulan data dan analisis statistik hasil klasifikasi. Analisis data dilakukan dengan menginterpretasikan hasil model terhadap tujuh kategori keluhan utama: masalah layanan, fitur aplikasi, pengalaman pengguna, harga paket mahal, kendala pembayaran, notifikasi, serta keamanan dan privasi.



Gambar 1. Tahapan Data Mining

1. Pengumpulan Data

Pengumpulan data ulasan pengguna dari platform seperti Google Play Store merupakan langkah awal yang umum dalam analisis sentimen aplikasi *mobile* [12].

	text
0	jaringan bab!..
1	Telkomsel Kartu Halo semakin buruk jaringannya...
2	fitur masih tetep aman dan nyaman tapi semenja...
3	ini kenapa jadi lemot banget gak bisa sama sek...
4	telkomsel super lelet.super delay.internet bur...
...	...
995	aplikasi nya sangat berguna untuk membeli pake...
996	makin lama makin lemot aja aplikasi ini. heran...
997	Baik
998	aplikasi mantab #Bergengsi
999	sang@t membantu

1000 rows × 1 columns

Gambar 2. Contoh Data *Crawling* dari myTelkomsel

Data keluhan untuk contoh ini dikumpulkan dari ulasan pengguna aplikasi myTelkomsel di Google Play Store dalam rentang waktu 1 Oktober 2024 – 28 Februari

2025. Data yang diambil berupa teks ulasan pengguna yang berisi keluhan terkait pengalaman mereka dalam menggunakan aplikasi.

Tabel 1. Sampel *dataset* hasil *Crawling*

no	***text***
1	Aplkasi tda Bejalan!
2	Sang@t membantu…
3	apkilasi mantab #Bergengsi

2. Katagori keluhan & *Labelling*

Keluhan pengguna diklasifikasikan ke dalam tujuh kategori utama berdasarkan analisis manual dan kajian literatur:

1. Masalah layanan: Keluhan yang berkaitan dengan gangguan jaringan, koneksi internet yang lambat, atau paket data yang tidak aktif setelah pembelian .
2. Fitur aplikasi: Masalah teknis dalam aplikasi, seperti kesulitan dalam *login*, *error* saat melakukan transaksi pembelian paket data, atau fitur yang tidak berfungsi dengan baik [5]
3. Pengalaman pengguna: Keluhan terkait tampilan dan performa aplikasi, seperti UI/UX yang kurang responsif, navigasi yang membingungkan, atau aplikasi yang sering mengalami *crash* [6].
4. Harga paket yang mahal: Keluhan mengenai perbedaan harga paket data di tiap nomor telepon, serta ketidaksesuaian antara harga yang diiklankan dengan harga yang berlaku [7]
5. Kendala pembayaran: Pengguna mengalami kegagalan dalam proses pembayaran menggunakan metode tertentu, seperti *e-wallet* atau kartu kredit yang tidak terverifikasi [5]
6. Notifikasi dan layanan pelanggan: Keluhan tentang kurangnya notifikasi yang relevan atau keterlambatan dalam respon layanan pelanggan terhadap permasalahan pengguna juga sering muncul [7].
7. Keamanan dan privasi: Pengguna mengeluhkan masalah keamanan akun, seperti *login* tidak sah, kebocoran data pribadi, atau kesulitan dalam mengganti kata sandi [8].

Labelling

Pada tahap ini, proses pelabelan data dilakukan menggunakan dua pendekatan, yaitu metode berbasis leksikon *(Lexicon-Based)* dan pelabelan secara manual. Metode *Lexicon-Based* melibatkan penggunaan daftar kata-kata yang telah memiliki label dan bobot tertentu, seperti yang ditunjukkan pada Gambar 2. Label-label ini disusun oleh peneliti sebelumnya dan mengklasifikasikan kata ke dalam beberapa label seperti aplikasi, fitur, pembayaran, promo, pelayanan, jaringan, notifikasi, dan *customer_service*. Penentuan klasifikasi keluhan sebuah kalimat dilakukan dengan menetapkan jenis-jenis dari keluhan pengguna aplikasi dan dimasukkan ke dalam label yang semestinya [20] [13].

itu, pelabelan manual dilakukan secara langsung oleh manusia dengan menilai konten dari data, seperti komentar atau ulasan, untuk menentukan ke arah mana keluhan itu disampaikan. Pendekatan ini sering digunakan untuk menghasilkan

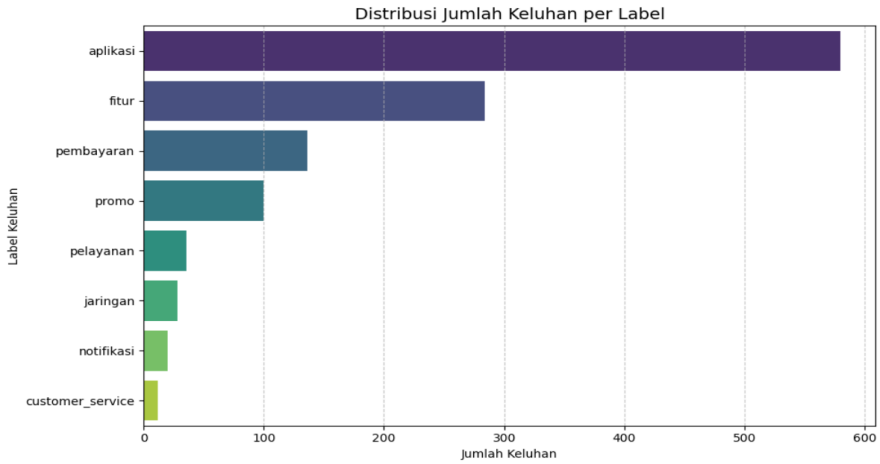
dataset dengan kualitas yang tinggi karena mempertimbangkan pemahaman dan intuisi manusia. Namun, proses ini memerlukan waktu, tenaga, dan sumber daya yang besar, terutama ketika jumlah data yang harus diberi label cukup banyak.

Setelah dilakukannya pelabelan maka hasilnya tertera seperti pada Gambar 4 dibawah ini.

```
                                          content             label
0          Sering log out sendiri tanpa alasan          aplikasi
1  Pembayaran berhasil tapi paket tidak masuk        pembayaran
2    Customer service sangat lambat responnya  customer_service
3    Customer service sangat lambat responnya  customer_service
4    Notifikasi terlalu sering dan mengganggu        notifikasi
label
aplikasi            580
fitur               284
pembayaran          136
promo               100
pelayanan            36
jaringan             28
notifikasi           20
customer_service     12
```

Gambar 3. Contoh *labelling*

Pada Gambar 4 dapat dilihat bahwa keluhan yang terhadap label aplikasi sangat banyak dengan nilai 520 keluhan.



Gambar 4. Contoh *Plot* dari hasil *labelling Lexicon*

Sementara keluhan lainnya seperti yang mengarah ke *customer_service*, notifikasi, jaringan dan sebagainya tidak sebanyak keluhan terhadap label aplikasi. Untuk mengetahui nilai keluhan tiap label dapat dilihat pada Gambar 3.

3. *Preprocessing Data* (Pemrosesan Awal)

Sebelum dilakukan klasifikasi, data teks yang diperoleh harus melalui serangkaian tahap pemrosesan untuk menghilangkan elemen yang tidak relevan dan mengubah teks menjadi format yang lebih terstruktur. Tahapan *preprocessing* meliputi:

1. ***Cleaning*** **:** Merupakan tahap dibersihkannya *dataset* dari elemen-elemen yang tidak berhubungan dengan informasi penting, seperti tautan URL, karakter khusus, angka, dan emotikon. Pada tahap ini juga telah dilakukan penghapusan komentar *duplicate* dan *missing data*, sehingga dari dataset awal 16.920 menjadi 15.202 komentar. Pada Tabel 2 adalah hasil dari pembersihan pada *dataset*.

Tabel 2. *Cleaning* pada *dataset*

No	***original_text***	***cleaned_text***
1	jaringan bab!	jaringan bab
2	Baik	Baik
3	apkilasi mantab #Bergengsi	aplikasi mantab Bergengsi

2. ***Case Folding*** **:** Tahap mengubah semua huruf dalam *dataset* menjadi huruf kecil, dilakukan agar menjaga konsistensi saat memproses *dataset*, seperti pada Tabel 3 dibawah ini.

Tabel 3. Hasil setelah melakukan *Case Folding* pada *dataset*

No	***cleaned_text***	***folded_text***
1	jaringan bab	jaringan bab
2	Baik	baik
3	aplikasi mantab Bergengsi	aplikasi mantab bergengsi

3. ***Normalization :*** Merupakan tahap untuk menyederhanakan teks dengan mengganti kata-kata yang tidak baku atau yang tidak sesuai dengan ejaan dalam kamus bahasa baku [15]. Pada Tabel 4 terlihat perubahan dari kata “mantab” yang merupakan kata lain dari “mantap”. Sehingga meningkatkan keterbacaan dan mudah untuk dianalisis.

Tabel 4. Setelah melakukan *Nomralization* pada *dataset*

no	***folded_text***	***normalized_text***
1	jaringan bab	jaringan babi
2	baik	baik
3	aplikasi mantab bergengsi	aplikasi mantap bergengsi

4. ***Stopword Removal*** **:** Menghapus kata-kata umum yang tidak memiliki makna signifikan, seperti "dan", "atau", "yang". Penghapusan *stopword* bertujuan untuk mengeliminasi kata-kata umum yang tidak memiliki makna signifikan dalam analisis teks [15].

Tabel 5. *Stopward Removing dataset*

no	***normalized_text***	***stopwords_removed_text***
1	jaringan babi	jaringan babi
2	baik	baik
3	aplikasi mantap bergengsi	aplikasi mantap bergengsi

5. **Tokenisasi :** Memisahkan teks menjadi kata-kata individu. Tokenisasi adalah proses memisahkan teks menjadi unit-unit kata atau token yang lebih kecil untuk memudahkan analisis selanjutnya [12].

Tabel 6. *Tokenizing* dataset

no	***stopwords_removed_text***	***tokenized_text***
1	jaringan babi	[jaringan, babi]
2	baik	[baik]
3	aplikasi mantap bergengsi	[aplikasi, mantap, bergengsi]

6. ***Stemming* :** Mengubah kata ke bentuk dasarnya, misalnya "berjalan" menjadi "jalan". *Stemming* mengubah kata ke bentuk dasarnya untuk menyamakan berbagai bentuk kata yang memiliki makna serupa [14].

Tabel 7. *Stemming*

no	***tokenized_text***	***stemmed_text***
1	[jaringan, babi]	jaringan babi
2	[baik]	baik
3	[aplikasi, mantap, bergengsi]	aplikasi mantap gengsi

4. Representasi Data (Pembobotan TF-IDF):

Mengubah teks ke dalam format numerik menggunakan metode TF-IDF (*Term Frequency-Inverse Document Frequency*) atau *Bag of Words*, sehingga bisa diolah oleh model *Naïve Bayes*.

Metode TF-IDF sering digunakan untuk mengubah teks menjadi format numerik yang dapat diolah oleh algoritma *machine learning* [15] [16].

$$TF(t,d) = \frac{t}{d}$$

t : frekuensi kata yang akan dihitung.

d : jumlah dokumen yang sedang dianalisis.

IDF adalah sebuah ukuran yang digunakan untuk menilai seberapa penting atau jarangnya suatu kata dalam suatu kumpulan dokumen.

$$IDF(t) = log\ log\ \left(\frac{d}{df(t)}\right)$$

$df(t)$: jumlah dokumen yang mengandung kata *t.*

d : jumlah dokumen yang sedang dianalisis.

Log : fungsi logaritma.

$$W(t,d) = TF(t,d) * IDF(t)$$

W : nilai tingkat kepentingan suatu kata dibandingkan dengan seluruh dokumen.

Klasifikasi:

Naïve Bayes merupakan metode klasifikasi yang menggunakan pendekatan probabilistik, didasarkan pada Teorema *Bayes* dengan asumsi bahwa setiap fitur bersifat independen satu sama lain (asumsi "*naive*"). Algoritma ini menghitung peluang untuk masing-masing kelas berdasarkan fitur *input*, kemudian memilih kelas dengan probabilitas terbesar sebagai hasil prediksi.

$$P(B) = \frac{P(A) * P(A)}{P(B)}$$

5. Pelatihan dan Pengujian Model

Pembagian *dataset* menjadi data latih dan data uji adalah praktik umum untuk mengevaluasi kinerja model secara objektif [17]. Setelah data siap, model *Naïve Bayes* dilatih menggunakan data yang telah diklasifikasikan sebelumnya. *Dataset* dibagi menjadi dua bagian:

- Data latih *(training data*): 80% dari total data digunakan untuk melatih model.
- Data uji (*testing data*): 20% dari total data digunakan untuk menguji performa model dan mengukur akurasi prediksi.

Dengan rumus:

$$Data\ Latih = \frac{80}{100} \times N$$

$$Data\ Uji = \frac{20}{100} \times N$$

Di mana:

- N = Jumlah total data dalam *dataset*
- Data Latih = Data yang digunakan untuk melatih model
- Data Uji = Data yang digunakan untuk menguji model

6. Evaluasi Kinerja Model

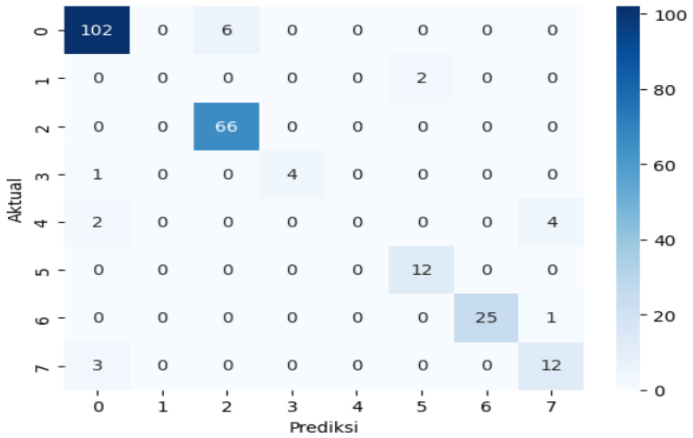
1. **Akurasi**: Akurasi mengukur seberapa banyak prediksi yang benar dibandingkan dengan total prediksi yang dibuat oleh model [18]. Akurasi adalah ukuran yang menunjukkan seberapa baik model klasifikasi dalam memprediksi kategori yang benar. Ini dihitung dengan membandingkan jumlah prediksi yang benar dengan total jumlah prediksi yang dibuat oleh model. Akurasi dinyatakan dalam bentuk persentase dan dihitung dengan rumus:

$$Akurasi = \frac{Jumlah\ Prediksi\ Benar}{Total\ Prediksi} \times 100\%$$

Di mana:

- Jumlah Prediksi Benar = Data uji yang diklasifikasikan dengan benar oleh model
- Total Prediksi = Seluruh data uji

Meskipun akurasi adalah metrik yang mudah dipahami, ia bisa menyesatkan, terutama dalam *dataset* yang tidak seimbang, di mana satu kategori mungkin jauh lebih dominan daripada yang lain.



Gambar 6. Prediksi dan Aktual

2. ***Precision***: *Precision* menunjukkan proporsi prediksi positif yang benar dari semua prediksi positif yang dibuat oleh model [18]. Precision, atau ketepatan, mengukur seberapa banyak dari semua prediksi positif yang benar-benar merupakan positif. Dalam konteks klasifikasi keluhan, precision menunjukkan proporsi keluhan yang benar-benar masuk dalam kategori tertentu dibandingkan dengan semua keluhan yang diklasifikasikan ke kategori tersebut. *Precision* dihitung dengan rumus:

$$Precision = \frac{TP}{TP + FP}$$

Di mana:

- TP (*True Positive*) = Jumlah prediksi benar untuk kelas positif,
- FP (*False Positive*) = Jumlah prediksi salah untuk kelas positif,
- FN (*False Negative*) = Jumlah prediksi salah untuk kelas negatif.

Precision yang tinggi menunjukkan bahwa model jarang salah dalam mengklasifikasikan keluhan ke dalam kategori yang salah, sehingga memberikan kepercayaan lebih pada hasil yang dikeluarkan.

3. ***Recall***: *Recall* mengukur kemampuan model dalam mendeteksi semua *instance* positif yang ada dalam *dataset* [18]. *Recall*, atau sensitivitas, mengukur seberapa banyak dari semua keluhan yang seharusnya diklasifikasikan ke

dalam kategori tertentu yang berhasil diklasifikasikan dengan benar oleh model. Ini menunjukkan kemampuan model untuk menangkap semua contoh positif dalam kategori tersebut. *Recall* dihitung dengan rumus:

$$Recall = \frac{TP}{TP + FN}$$

Di mana:

- TP (*True Positive*) = Jumlah prediksi benar untuk kelas positif,
- FP (*False Positive*) = Jumlah prediksi salah untuk kelas positif,
- FN (*False Negative*) = Jumlah prediksi salah untuk kelas negatif.

Recall yang tinggi berarti model mampu mengidentifikasi sebagian besar keluhan yang relevan, tetapi mungkin juga mengorbankan *precision*, karena bisa jadi model mengklasifikasikan beberapa keluhan yang tidak relevan sebagai positif.

4. ***F1-Score***: *F1-Score* adalah rata-rata harmonis dari *precision* dan *recall*, memberikan keseimbangan antara keduanya [24]. *F1-Score* adalah metrik yang menggabungkan *precision* dan *recall* menjadi satu nilai tunggal, memberikan gambaran yang lebih seimbang tentang kinerja model. *F1-Score* dihitung sebagai rata-rata harmonis dari *precision* dan *recall*, dan sangat berguna ketika kita ingin menemukan keseimbangan antara keduanya. *F1-Score* dihitung dengan rumus:

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

F1-Score yang tinggi menunjukkan bahwa model memiliki baik *precision* maupun *recall* yang baik, sehingga dapat diandalkan dalam klasifikasi.

5. ***Confusion Matrix***: *Confusion Matrix* adalah tabel yang digunakan untuk menggambarkan kinerja model klasifikasi dengan menunjukkan jumlah prediksi yang benar dan salah untuk setiap kategori. *Confusion matrix* memberikan gambaran lengkap tentang kinerja model dengan menunjukkan jumlah prediksi benar dan salah untuk setiap kelas [18]. Tabel ini terdiri dari empat komponen utama:

Actual \ *Predicted*	Prediksi Positif	Prediksi Negatif
Aktual Positif	*True Positives (TP)*	*False Negatives (FN)*
Aktual Negatif	*False Positives (FP)*	*True Negatives (TN)*

Confusion Matrix memiliki wawasan yang lebih mendalam tentang bagaimana model membedakan antara kategori yang berbeda dan membantu dalam mengidentifikasi kesalahan klasifikasi. Dengan menganalisis *confusion matrix*, kita dapat memahami di mana model melakukan kesalahan dan bagaimana cara memperbaikinya.

3. Hasil Dan Pembahasan

Data dalam penelitian ini diperoleh langsung dari ulasan pengguna aplikasi MyTelkomsel yang terdapat di platform Google Play Store. Pengambilan data dilakukan pada periode 1 Mei 2025 hingga 1 September 2025, dengan tujuan memperoleh representasi yang relevan terhadap pengalaman pengguna dalam rentang waktu terbaru.

Proses pengambilan data dilakukan oleh penulis, bahasa pemrograman Python adalah alat yang dijalankan pada lingkungan Google Colab, dengan memanfaatkan teknik web crawling untuk mengekstraksi teks ulasan beserta informasi pendukung seperti tanggal unggahan dan rating pengguna. Tahap pengumpulan data ini bertujuan untuk memperoleh kumpulan ulasan yang autentik dan bersumber langsung dari pengalaman pengguna aplikasi. Data yang diambil difokuskan pada ulasan yang mengandung keluhan atau umpan balik negatif, karena penelitian ini berfokus pada klasifikasi keluhan pengguna. Setelah data berhasil dikumpulkan, dilakukan tahap penyaringan awal untuk menghapus data yang bersifat duplikat, kosong (missing values), atau tidak relevan, seperti komentar promosi dan ulasan yang tidak berkaitan dengan fungsi aplikasi.

Hasil dari proses pengumpulan ini menghasilkan 800 data ulasan pengguna yang kemudian digunakan sebagai dataset penelitian. Dataset ini menjadi dasar bagi tahap selanjutnya, yaitu preprocessing teks, pelabelan (labelling), serta proses training dan testing model Naïve Bayes. Melalui pengumpulan data secara langsung dari Google Play Store, penelitian ini memastikan bahwa data yang digunakan bersifat aktual, kontekstual, dan mencerminkan persepsi nyata pengguna terhadap kinerja aplikasi MyTelkomsel.

Data Labelling

Pada Gambar 7 ditunjukkan contoh hasil proses pelabelan data ulasan pengguna aplikasi *myTelkomsel* yang dilakukan menggunakan pendekatan *lexicon-based*.

Setiap ulasan (review) diklasifikasikan ke dalam salah satu dari delapan kategori keluhan, masalah layanan, fitur, pembayaran, harga paket, pengalaman pengguna, jaringan, notifikasi, dan privasi, berdasarkan kata kunci yang muncul dalam teks. Sebagai contoh, ulasan yang mengandung kata *"error"* atau *"tidak bisa dibuka"* diberi label aplikasi, sedangkan ulasan dengan kata *"pembayaran gagal"* atau *"saldo tidak masuk"* dikategorikan sebagai pembayaran. Proses ini dilakukan secara otomatis menggunakan kamus kata kunci (lexicon) yang telah disusun berdasarkan kajian literatur dan analisis konteks keluhan pengguna.

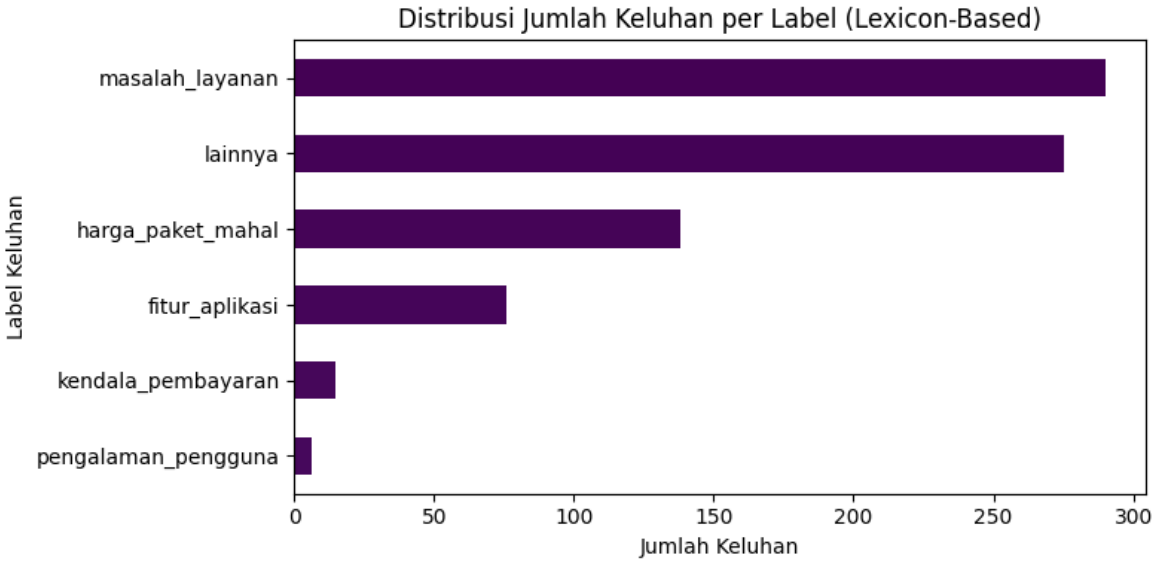
```
=== Persentase Tiap Kategori Keluhan ===
label
masalah_layanan        36.25
lainnya                34.38
harga_paket_mahal      17.25
fitur_aplikasi          9.50
kendala_pembayaran      1.88
pengalaman_pengguna     0.75
```

content	score	label
Mohon iktikad baik dari pihak telkomsel terkait fitur voucher saldo darurat, dikare	1	fitur_aplikasi
Halo mu Telkomsel,Atas mohon maaf nya ini daily check harian kok nga bisa kebu	4	lainnya
kuota dari daily check-in gak bisa dipake padahal udh effort buat login setiap hari,	1	masalah_layanan
ini kenapa aplikasi tidak bisa beli kuota?? selalu muncul tulisan , cek koneksi and	2	harga_paket_mahal
awesome	3	lainnya
buruk sekali sekarang tay . paket kek tay naik turun seenaknya tay . bikin masyara	2	harga_paket_mahal
sangat payah, susah untuk di gunakan.	1	lainnya
4G nya di tempat Saya kurang	5	lainnya
saya sangat puas dengan layanan Telkomsel	5	lainnya
mantab	5	lainnya
tolong diperbaiki pembayaran via apk dana..	1	kendala_pembayaran
kenapa susah banged mau beli paket, keluar keluar terus	1	harga_paket_mahal
harga makin mahal , paket yang didapat makin sedikit	1	harga_paket_mahal
aplikasi yang bagus	5	lainnya
sering banget banyak malingnya ni provider, masa ilsi pulsa 10 rb blom dipaketin	1	harga_paket_mahal
perbaiki jaringan baru kuberi bintang 5 !!!	1	masalah_layanan
kartu tenggang, diisi pulsa nggak masuk padahal udah jelas transaksi berhasil, gili	1	fitur_aplikasi
sangat membantu	5	lainnya
sering" promonya dong 😊🙏	5	lainnya
tsel MALING MALING MALING PULSAAAA wooyy MALING MALING MALING MALIN	1	lainnya

Gambar 7. Hasil presentase *Labelling* dan hasil *Labelling*

Sementara pada Gambar 8 memperlihatkan hasil visualisasi jumlah keluhan pada setiap kategori berdasarkan proses pelabelan sebelumnya. Terlihat bahwa keluhan dengan label "masalah_layanan" memiliki jumlah paling banyak, yaitu sekitar *290 data ulasan*, diikuti oleh lainnya, harga_paket_mahal, fitur_aplikasi, kendala_pembayaran dan pengalaman_pengguna. Hal ini menunjukkan bahwa permasalahan yang paling sering dialami pengguna *myTelkomsel* berkaitan dengan masalah layanan, seperti error, gagal login, atau aplikasi tidak merespons. Sementara itu, keluhan terkait harga palet, fitur aplikasi, kendala pembayaran dan pengalaman pengguna memiliki frekuensi yang lebih rendah, menandakan bahwa masalah pada aspek tersebut relatif lebih jarang dikeluhkan oleh pengguna.

Distribusi keluhan ini memberikan gambaran awal tentang area layanan yang paling perlu diperbaiki oleh Telkomsel, serta menjadi dasar bagi proses *training* model *Naïve Bayes* untuk klasifikasi otomatis pada tahap selanjutnya.



Gambar 8. *Plot* dari hasil *labelling Lexicon*

Preprocessing Data

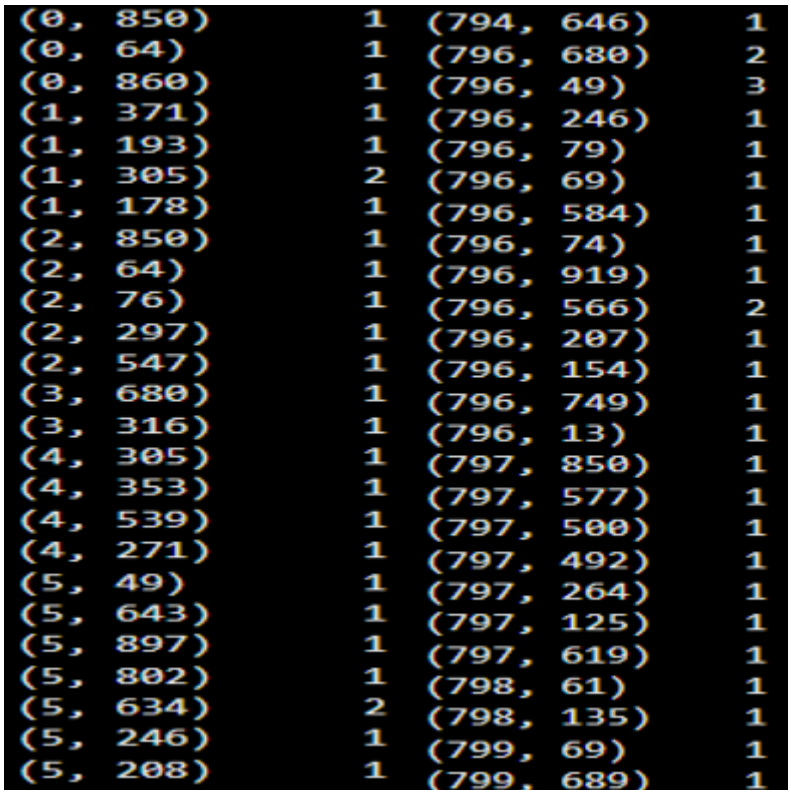
Pada tahap preprocessing, data ulasan pengguna aplikasi MyTelkomsel diproses melalui beberapa langkah utama, yaitu *cleaning, case folding, normalization, stopword removal, tokenisasi,* dan *stemming*. Tahap ini dilakukan agar data teks yang semula tidak terstruktur dapat diubah menjadi bentuk yang lebih bersih dan siap digunakan pada tahap representasi data dengan metode TF-IDF.

content	cleaned	normalized	stop_removed	tokenized	stemmed
sinyal terbaik end stabil	sinyal terbaik end stabil	sinyal terbaik end stabil	sinyal terbaik end stabil	['sinyal', 'terbaik', 'end', 'stabil']	sinyal baik end stabil
kartu elit jaringan sulit,hujan sikit jaringan down	kartu elit jaringan sulithujan sikit jaringan down	kartu elit jaringan sulithujan sikit jaringan down	kartu elit jaringan sulithujan sikit jaringan down	['kartu', 'elit', 'jaringan', 'sulithujan', 'sikit', 'jaringan', 'down']	kartu elit jaring sulithujan sikit jaring down
tolong perbarui sinyalnya menjadi lebih baik	tolong perbarui sinyalnya menjadi lebih baik	tolong perbarui sinyalnya menjadi lebih baik	perbarui sinyalnya menjadi lebih baik	['perbarui', 'sinyalnya', 'menjadi', 'lebih', 'baik']	baru sinyal jadi lebih baik
internett nya jelek bgttt ajgg	internett nya jelek bgttt ajgg	internett nya jelek bgttt ajgg	internett nya jelek bgttt ajgg	['internett', 'nya', 'jelek', 'bgttt', 'ajgg']	internett nya jelek bgttt ajgg
kalau pln padam langsung jaringan hilang.	kalau pln padam langsung jaringan hilang	kalau pln padam langsung jaringan hilang	kalau pln padam langsung jaringan hilang	['kalau', 'pln', 'padam', 'langsung', 'jaringan', 'hilang']	kalau pln padam langsung jaring hilang
Aplikasi My Telkomsel sangat mudah digunakan, fiturnya lengkap, dan mudah untuk semua kalangan dalam hal bertransaksi.	aplikasi my telkomsel sangat mudah digunakan fiturnya lengkap dan mudah untuk semua kalangan dalam hal bertransaksi	aplikasi my telkomsel sangat mudah digunakan fiturnya lengkap dan mudah untuk semua kalangan dalam hal bertransaksi	aplikasi my telkomsel sangat mudah digunakan fiturnya lengkap mudah semua kalangan hal bertransaksi	['aplikasi', 'my', 'telkomsel', 'sangat', 'mudah', 'digunakan', 'fiturnya', 'lengkap', 'mudah', 'semua', 'kalangan', 'hal', 'bertransaksi']	aplikasi my telkomsel sangat mudah guna fiturnya lengkap mudah semua kalang hal transaksi

Gambar 9. Hasil Preprocessing

TF-IDF

Pada Gambar 10 ditunjukkan proses pembagian data (data split) sebelum dilakukan klasifikasi menggunakan algoritma Naïve Bayes. Data ulasan dibagi menjadi dua bagian, yaitu 80% sebagai data training dan 20% sebagai data testing. Data training digunakan untuk melatih model, sedangkan data testing digunakan untuk menguji performa model pada data yang belum pernah dipelajari sebelumnya. Teknik pembagian data ini umum digunakan dalam penelitian klasifikasi teks untuk memperoleh hasil evaluasi yang lebih objektif.



(0, 850)	1	(794, 646)	1
(0, 64)	1	(796, 680)	2
(0, 860)	1	(796, 49)	3
(1, 371)	1	(796, 246)	1
(1, 193)	1	(796, 79)	1
(1, 305)	2	(796, 69)	1
(1, 178)	1	(796, 584)	1
(2, 850)	1	(796, 74)	1
(2, 64)	1	(796, 919)	1
(2, 76)	1	(796, 566)	2
(2, 297)	1	(796, 207)	1
(2, 547)	1	(796, 154)	1
(3, 680)	1	(796, 749)	1
(3, 316)	1	(796, 13)	1
(4, 305)	1	(797, 850)	1
(4, 353)	1	(797, 577)	1
(4, 539)	1	(797, 500)	1
(4, 271)	1	(797, 492)	1
(5, 49)	1	(797, 264)	1
(5, 643)	1	(797, 125)	1
(5, 897)	1	(797, 619)	1
(5, 802)	1	(798, 61)	1
(5, 634)	2	(798, 135)	1
(5, 246)	1	(799, 69)	1
(5, 208)	1	(799, 689)	1

Gambar 10. Fitur Ekstrasi menggunakan *Count Vectorizer*

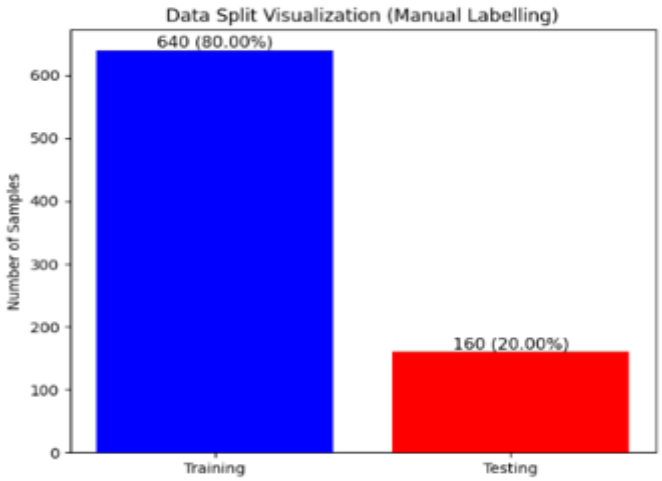
```
Komentar asli: aplikasi my telkomsel sangat mudah guna fiturnya lengkap mudah semua kalang hal transaksi

<Compressed Sparse Row sparse matrix of dtype 'float64'
        with 12 stored elements and shape (1, 2015)>
  Coords        Values
  (0, 97)       0.13890392382190112
  (0, 1187)     0.2343865198625832
  (0, 1789)     0.13819542561511095
  (0, 1570)     0.18953508093240679
  (0, 1174)     0.4950773073393561
  (0, 586)      0.20926477683524836
  (0, 500)      0.3099647244500355
  (0, 997)      0.33602247632316085
  (0, 1617)     0.25392335712036107
  (0, 787)      0.35670567757371624
  (0, 615)      0.33602247632316085
  (0, 1890)     0.2612230815330342
(5, 97) 0.13890392382190112
(5, 1187) 0.2343865198625832
(5, 1789) 0.13819542561511095
(5, 1570) 0.18953508093240679
(5, 1174) 0.4950773073393561
(5, 586) 0.20926477683524836
(5, 500) 0.3099647244500355
(5, 997) 0.33602247632316085
(5, 1617) 0.25392335712036107
(5, 787) 0.35670567757371624
(5, 615) 0.33602247632316085
(5, 1890) 0.2612230815330342
```

Gambar 11. Hasil dari TF-IDF dari komentar indeks ke 5

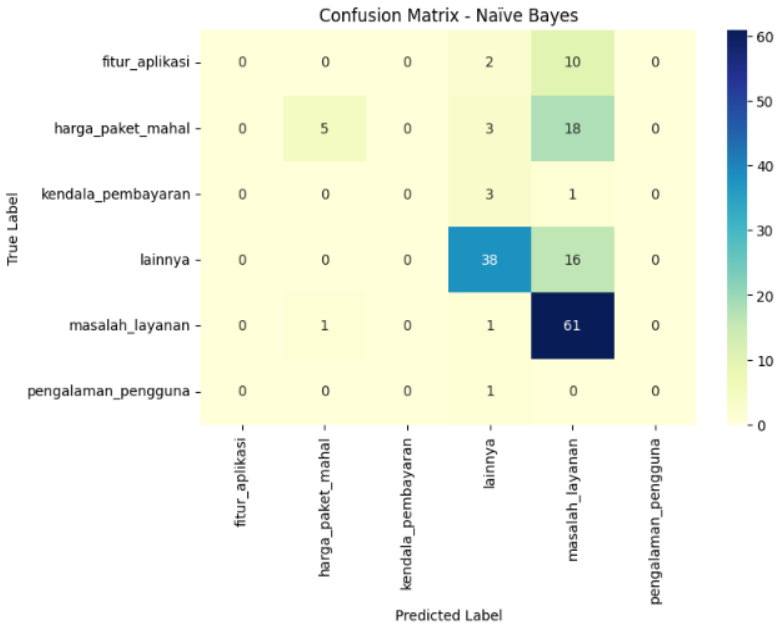
Analisis Klasifikasi dan Evaluasi

Pada tahap ini dilakukan proses klasifikasi menggunakan algoritma Naïve Bayes terhadap data ulasan yang telah dibagi sebelumnya menjadi data training dan testing dengan komposisi 80% dan 20%. Evaluasi dilakukan dengan menggunakan confusion matrix untuk mengetahui performa model dalam mengklasifikasikan ulasan positif dan negatif. Hasil visualisasi confusion matrix dapat dilihat pada Gambar 12 di bawah ini.

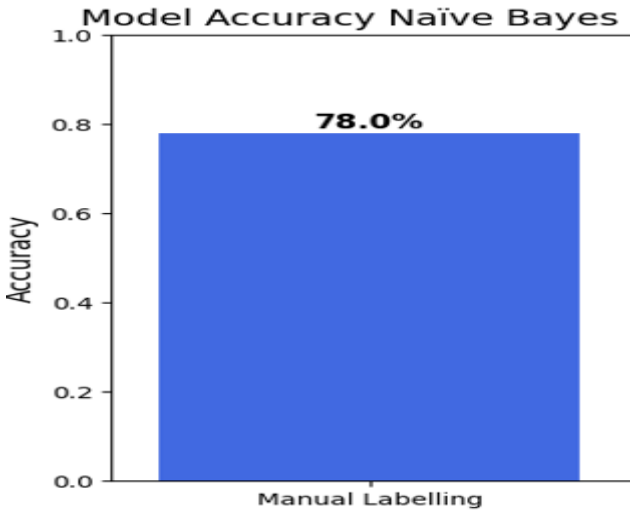


Gambar 12. Data split

Berdasarkan hasil confusion matrix pada Gambar 13, algoritma Naïve Bayes dengan metode manual labelling menghasilkan nilai akurasi sebesar 53,8%, dengan nilai precision 54,5%, recall 71,4%, dan f1-score 61,8% pada kelas positif.



Gambar 13. *Confusion matrix* Naïve Bayes



Gambar 14. *Model Accuracy* Naïve Bayes

Dari hasil pengujian yang ditunjukkan pada Gambar 14, algoritma Naïve Bayes menghasilkan tingkat akurasi sebesar 78% pada data dengan pelabelan manual. Secara keseluruhan, hasil ini memperlihatkan bahwa algoritma Naïve Bayes mampu melakukan klasifikasi dengan baik pada pendekatan manual.

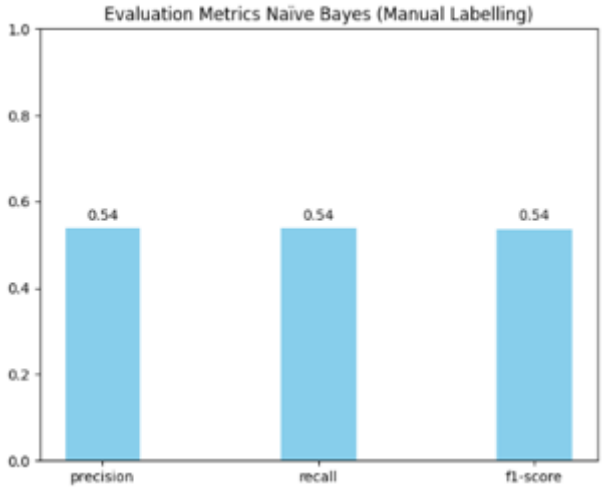
```
=== Classification Report ===
                    precision    recall  f1-score   support

      fitur_aplikasi       0.00      0.00      0.00        12
   harga_paket_mahal       0.83      0.19      0.31        26
  kendala_pembayaran       0.00      0.00      0.00         4
             lainnya       0.79      0.70      0.75        54
     masalah_layanan       0.58      0.97      0.72        63
pengalaman_pengguna       0.00      0.00      0.00         1

            accuracy                           0.65       160
           macro avg       0.37      0.31      0.30       160
        weighted avg       0.63      0.65      0.59       160
```

Gambar 15. *Classification Report* Naïve Bayes

Seperti yang terlihat pada Gambar 15, hasil *classification report* memberikan gambaran detail mengenai nilai *precision, recall,* dan *f1-score* dari algoritma Naïve Bayes dengan pelabelan manual. Pada Gambar 16 disajikan visualisasi evaluasi metrik yang sama dalam bentuk diagram batang.



Gambar 16. *Evaluation Metrics by Model and Class*

Pada pelabelan manual, algoritma Naïve Bayes menunjukkan hasil evaluasi dengan nilai precision sebesar 51%, artinya hanya sekitar setengah dari prediksi positif yang dibuat model benar-benar sesuai dengan label aslinya. Kemampuan model dalam mengenali data positif (recall) tercatat sebesar 38%, sehingga nilai f1-score yang dihasilkan juga rendah, yaitu 43%. Hasil ini menggambarkan bahwa model masih kesulitan dalam mengklasifikasikan ulasan positif secara akurat pada pelabelan manual.

Validasi Data

Visualisasi data dalam penelitian ini memberikan dan menyajikan informasi dalam bentuk *word cloud*. Dimana hasil visualisasi ini dibagi menjadi 2 bagian yaitu, *word cloud* positif berwarna hijau dan negatif berwarna oranye dari data dengan pelabelan secara manual dan *lexicon*. Pada Gambar 17 ditampilkan hasil analisis 10 kata terbanyak (*Most Frequent Words*) yang muncul pada dataset ulasan aplikasi *MyTelkomsel*. Visualisasi ini menunjukkan bahwa kata-kata seperti “telkomsel”, “paket”, “nya”, “mahal”, “beli”, “aplikasi”, “jaringan”, “kuota”, “gak” dan “ga” mendominasi percakapan dalam ulasan pengguna.

```
=== 10 Kata Paling Sering Muncul ===
telkomsel : 176
paket : 171
nya : 157
beli : 134
mahal : 130
aplikasi : 125
jaringan : 125
kuota : 107
gak : 95
ga : 94
```

Gambar 17. *Most Frequent Words*

Gambar 18. *Word Cloud Manual Labelling*

Berdasarkan Gambar 18 di atas, hasil Word Cloud tersebut terlihat bahwa kata seperti "telkomsel", "aplikasi", "jaringan", "pulsa", "lemot", "kuota", dan "sinyal" muncul dengan ukuran yang lebih besar, menandakan bahwa istilah-istilah tersebut paling sering digunakan dalam ulasan negatif pengguna. Hal ini menunjukkan bahwa sebagian besar keluhan berhubungan dengan kualitas jaringan, performa aplikasi yang lambat, serta permasalahan terkait kuota dan pulsa. Dengan demikian, Word Cloud pada Gambar 18 memberikan gambaran umum mengenai fokus utama keluhan pengguna sebelum dilakukan proses klasifikasi menggunakan metode Naïve Bayes.

Model Naïve Bayes menunjukkan akurasi sebesar 78%, dengan precision 54,5%, recall 71,4%, dan F1-score 61,8%. Hasil ini sejalan dengan penelitian Rahmawati & Santoso (2023) yang menggunakan Naïve Bayes untuk ulasan e-commerce Tokopedia dengan akurasi di atas 70%. Namun, dibandingkan studi Siregar (2024) pada aplikasi DANA yang menggunakan Support Vector Machine, model Naïve Bayes dalam penelitian ini memiliki efisiensi komputasi lebih tinggi. Kebaruan penelitian ini terletak pada penerapan klasifikasi berbasis keluhan (bukan sentimen) yang mampu mengidentifikasi masalah layanan paling dominan, yakni gangguan jaringan (36,25%). Hal ini menegaskan bahwa Naïve Bayes dapat diadaptasi tidak hanya untuk analisis sentimen umum tetapi juga untuk pemetaan masalah pengguna secara spesifik.

4. Simpulan

Melalui hasil penelitian yang sudah dilakukan oleh penulis, penerapan metode Naïve Bayes dalam mengklasifikasikan keluhan pengguna terhadap aplikasi MyTelkomsel terbukti mampu mengidentifikasi pola keluhan yang paling sering disampaikan oleh pengguna. Data yang digunakan sebanyak 800 ulasan negatif yang diambil dari Google Play Store periode 1 Mei hingga 1 September 2025 dan telah melalui proses pelabelan manual. Proses preprocessing teks yang meliputi tahapan cleaning, case folding, stopword removal, tokenisasi, dan stemming, serta pembobotan menggunakan TF-IDF berhasil mempersiapkan data dengan baik sebelum dilakukan klasifikasi.

Hasil pengujian menunjukkan bahwa algoritma Naïve Bayes menghasilkan nilai akurasi sebesar 78%, precision 54,5%, recall 71,4%, dan F1-score 61,8%. Nilai

tersebut menunjukkan bahwa model mampu mengenali dan mengelompokkan keluhan pengguna dengan tingkat ketepatan yang cukup baik. Berdasarkan hasil klasifikasi, kategori masalah layanan menjadi jenis keluhan yang paling dominan dengan persentase sebesar 36,25%, diikuti oleh kategori harga paket mahal (17,25%), fitur aplikasi (9,50%), kendala pembayaran (1,88%), dan pengalaman pengguna (0,75%). Temuan ini menunjukkan bahwa sebagian besar pengguna mengalami kendala pada koneksi jaringan, kestabilan aplikasi, serta harga paket yang dianggap tinggi.

Dengan demikian, dapat disimpulkan bahwa metode Naïve Bayes efektif digunakan dalam proses klasifikasi teks untuk menganalisis keluhan pengguna aplikasi MyTelkomsel. Hasil penelitian ini dapat menjadi acuan bagi pengembang dan pihak Telkomsel dalam memahami fokus utama permasalahan pengguna, sehingga dapat digunakan sebagai dasar untuk meningkatkan kualitas jaringan, memperbaiki performa aplikasi, serta meninjau kembali strategi harga paket agar lebih sesuai dengan kebutuhan pelanggan.

5. Daftar Pustaka

[1] H. -, A. Y. Kuntoro, and T. Asra, "Klasifikasi Keluhan Pengguna Kai Access Untuk Pemesanan Tiket Dengan Algoritma Svm Dan Naïve Bayes," *JIKA (Jurnal Informatika)*, vol. 6, no. 2, p. 161, 2022, doi: 10.31000/jika.v6i2.6187.

[2] M. D. K. Rizky Ade Safitri, Yoseph Tajul Arifin, "Kepuasan Layanan Aplikasi Depok Single Window," vol. 8, no. 3, pp. 4094–4102, 2024.

[3] N. A. Maulana and Z. Fatah, "Gudang Jurnal Multidisiplin Ilmu Penerapan Metode Naïve Bayes untuk Analisis Sentimen Ulasan Produk di Platform E-Commerce," vol. 2, no. November, pp. 433–439, 2024.

[4] L. Rahmawati and D. B. Santoso, "Implementasi Metode Naive Bayes Untuk Klasifikasi Ulasan Aplikasi E-Commerce Tokopedia," *INTECOMS: Journal of Information Technology and Computer Science*, vol. 6, no. 1, pp. 116–124, 2023, doi: 10.31539/intecoms.v6i1.5515.

[5] N. Devi *et al.*, "Pengaruh E-Service Quality Terhadap Kepuasan Pelanggan Pengguna Aplikasi Mytelkomsel Tahun 2024," vol. 10, no. 5, pp. 1104–1107, 2024.

[6] R. P. Sahartira and F. N. Hasan, "MYTELKOMSEL PADA GOOGLE PLAYSTORE MENGGUNAKAN," pp. 1–10, 2025.

[7] L. Rahmawati and D. B. Santoso, "Implementasi Metode Naive Bayes Untuk Klasifikasi Ulasan Aplikasi E-Commerce Tokopedia," *INTECOMS: Journal of Information Technology and Computer Science*, vol. 6, no. 1, pp. 116–124, 2023, doi: 10.31539/intecoms.v6i1.5515.

[8] A. Ibrahim, F. S. Elisa, J. Fernando, L. Salsabila, N. Anggraini, and S. N. Arafah, "Pengaruh E-Service Quality Terhadap Loyalitas Pengguna Aplikasi MyTelkomsel," *Building of Informatics, Technology and Science (BITS)*, vol. 3, no. 3, pp. 302–311, 2021, doi: 10.47065/bits.v3i3.1076.

[9] F. A. F. SIREGAR, "ANALISIS SENTIMEN BERBASIS ASPEK PADA ULASAN APLIKASI MYTELKOMSEL MENGGUNAKAN ALGORITMA NAÏVE BAYES DAN METODE TF-IDF," *Αγαη*, vol. 15, no. 1, pp. 37–48, 2024.

[10] S. A. Saragih and Z. Siregar, "Pengaruh Fitur Layanan , Privasi Keamanan dan Kepercayaan Terhadap Kepuasan Pengguna Aplikasi Dana Sebagai Alat Pembayaran (Studi Pada Mahasiswa Prodi Manajemen Unimed)," vol. 5, no. 1, pp. 129–140, 2025.

[11] R. Setiyawan and Z. Mustofa, "Comparison of the performance of naive bayes and support vector machine in sirekap sentiment analysis with the lexicon-based approach," *Journal of Soft Computing Exploration*, vol. 5, no. 2, pp. 122–132, Jun. 2024, doi: 10.52465/joscex.v5i2.367.

[12] D. Y. Ramadhan, *Analisis sentimen opini keamanan kode one time password dompet digital dalam media sosial menggunakan metode lexicon based dan support vector machine learning*, vol. 15, no. 1. 2024.

[13] E. S. Romaito, M. K. Anam, R. Rahmaddeni, and A. N. Ulfah, "Perbandingan Algoritma SVM Dan NBC Dalam Analisa Sentimen Pilkada Pada Twitter," *CSRID Journal*, vol. 13, no. 3, pp. 169–179, 2021.

[14] A. A. Syam, G. H. M, A. Salim, D. F. Surianto, and M. F. B, "Analisis teknik preprocessing pada sentimen masyarakat terkait konflik israel-palestina menggunakan support vector machine," vol. 9, no. 3, pp. 1464–1472, 2024.

[15] B. Hakim, "Analisa Sentimen Data Text Preprocessing Pada Data Mining Dengan Menggunakan Machine Learning," *JBASE - Journal of Business and Audit Information Systems*, vol. 4, no. 2, 2021, doi: 10.30813/jbase.v4i2.3000.

[16] D. Rifaldi, Abdul Fadlil, and Herman, "Teknik Preprocessing Pada Text Mining Menggunakan Data Tweet 'Mental Health,'" *Decode: Jurnal Pendidikan Teknologi Informasi*, vol. 3, no. 2, pp. 161–171, 2023, doi: 10.51454/decode.v3i2.131.

[17] G. Noer, "Implementasi Algoritma Naïve Bayes dan TF-IDF Dalam Analisis Sentimen Data Ulasan (Studi Kasus: Ulasan Review Aplikasi E-commerce Shopee di Situs Google …," *Repository.Uinjkt.Ac.Id*, 2023.

[18] H. Hajaroh, T. Suprapti, and R. Narasati, "Implementasi Algoritma Naive Bayes Untuk Analisis Sentimen Ulasan Produk Makanan Dan Minuman Di Tokopedia," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 1, pp. 111–118, 2024, doi: 10.36040/jati.v8i1.8237.

[19] T. Hartati, R. T. Sohadi, E. Tohidi, and E. Wahyudin, "Penerapan Algoritma Naive Bayes pada Analisis Sentimen Ulasan Aplikasi Whoosh – Kereta Cepat Di Google Play Store," vol. 6, no. 1, pp. 244–249, 2024.

[20] S. Budi, "Implementasi Metode Naive Bayes Untuk Klasifikasi Ulasan Pada Aplikasi Telegram," *INTECOMS: Journal of Information Technology and Computer Science*, vol. 6, no. 2, pp. 1282–1288, 2023, doi: 10.31539/intecoms.v6i2.8284.