Journal of Information Technology and Computer Science (INTECOMS)
Volume 9 Nomor 4, Tahun 2026
e-ISSN : 2614-1574
p-ISSN : 2621-3249

ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI *THREADS* PADA *GOOGLE PLAY STORE* DENGAN MENGGUNAKAN ALGORITMA *NAIVE BAYES*

SENTIMENT ANALYSIS OF USER REVIEWS OF THE THREADS APP ON THE GOOGLE PLAY STORE USING THE NAIVE BAYES ALGORITHM

Gita Tiara Aftita[1], Raissa Amanda Putri[2]
Ilmu Komputer, Fakultas Sains dan Teknologi, Universitas Islam Negeri Sumatera Utara[1,2]
gitatiara933@gmail.com[1]

ABSTRACT

User reviews on the Google Play Store can be used to gauge user feedback on the Threads app. This study aims to analyze the sentiment of user reviews for the Threads app on the Google Play Store using the Naive Bayes algorithm. A total of 5,000 reviews were collected using the Google Play Scraper and subsequently underwent text preprocessing, which included cleaning, case folding, normalization, tokenization, stopword removal, and stemming. After preprocessing, 3,312 reviews were obtained and then labeled with sentiment based on user ratings. Next, Term Frequency-Inverse Document Frequency (TF-IDF) vectorization was performed, and the data was split into training and test sets at an 80:20 ratio. Class imbalance in the training data was addressed through oversampling before training with the Naive Bayes algorithm. Model evaluation was conducted using a confusion matrix. Test results showed that the model achieved an accuracy of 70%. The best performance was achieved for positive sentiment with an F1-score of 0.83, while neutral sentiment had an F1-score of 0.24. These results indicate that the Naive Bayes model is better at recognizing patterns of positive sentiment, whereas neutral sentiment remains the most difficult class to identify even after oversampling has been applied.

Keywords*: Sentiment Analysis, Threads, Google Play Store, TF-IDF, Naive Bayes*

ABSTRAK

Ulasan pengguna pada *Google Play Store* dapat digunakan untuk mengetahui tanggapan pengguna terhadap aplikasi Threads. Penelitian ini bertujuan untuk menganalisis sentimen ulasan pengguna aplikasi Threads pada *Google Play Store* menggunakan algoritma *Naive Bayes*. Sebanyak 5.000 ulasan dikumpulkan menggunakan *Google Play Scraper* dan selanjutnya melalui tahap *text preprocessing* yang meliputi *cleaning*, *case folding*, normalisasi, tokenisasi, *stopword removal*, dan *stemming*. Setelah *preprocessing*, diperoleh 3.312 ulasan yang kemudian diberi label sentimen berdasarkan *rating* pengguna. Selanjutnya, dilakukan vektorisasi *Term Frequency-Inverse Document Frequency* (TF-IDF) dan dibagi menjadi data latih dan data uji dengan rasio 80:20. Ketidakseimbangan kelas pada data latih ditangani melalui *oversampling* sebelum proses pelatihan menggunakan algoritma Naive Bayes. Evaluasi model dilakukan menggunakan *confusion matrix*. Hasil pengujian menunjukkan bahwa model memperoleh *accuracy* sebesar 70%. Performa terbaik diperoleh pada sentimen positif dengan F1-*score* sebesar 0,83, sedangkan sentimen netral memiliki F1-*score* sebesar 0,24. Hasil tersebut menunjukkan bahwa *Naive Bayes* mampu mengenali pola sentimen positif dengan lebih baik, sedangkan sentimen netral masih menjadi kelas yang paling sulit dikenali meskipun *oversampling* telah diterapkan.

Kata Kunci: Analisis Sentimen, *Threads*, *Google Play Store*, TF-IDF, *Naive Bayes*

PENDAHULUAN

Perkembangan teknologi informasi telah mendorong perubahan dalam berbagai aktivitas masyarakat Indonesia, termasuk dalam memperoleh informasi, berkomunikasi, dan menggunakan layanan digital. Peningkatan akses terhadap internet menjadi salah satu indikator berkembangnya pemanfaatan teknologi digital di Indonesia. Berdasarkan data Badan Pusat Statistik (2025), sebanyak 72,78% penduduk Indonesia telah mengakses internet pada tahun 2024, meningkat dibandingkan 69,21% pada tahun 2023. Kondisi tersebut menunjukkan bahwa penggunaan internet telah menjadi bagian dari aktivitas masyarakat dan turut mendukung perkembangan berbagai layanan berbasis digital.

Berdasarkan survei Asosiasi Penyelenggara Jasa Internet Indonesia (2026), tingkat penetrasi internet Indonesia pada tahun 2025 mencapai 80,66% atau sekitar 229 juta pengguna. Angka tersebut

kembali meningkat pada tahun 2026 menjadi 81,72% dengan jumlah pengguna mencapai 235,26 juta orang. Peningkatan tersebut menunjukkan bahwa internet telah menjadi bagian yang semakin luas dalam aktivitas masyarakat Indonesia dan turut mendorong perkembangan berbagai layanan serta *platform* digital.

Kondisi ini menegaskan bahwa ruang digital kini menjadi arena penting dalam pembentukan partisipasi kewargaan, di mana generasi muda tidak hanya menjadi konsumen informasi tetapi juga aktor aktif dalam diskursus sosial dan politik di dunia maya (Ramdani et al., 2026). Media sosial merujuk pada berbagai teknologi yang memungkinkan kolaborasi, berbagi informasi, dan interaksi melalui pesan daring (Yusup, A et al., 2023). Seiring dengan perkembangan internet, teknologi dan fitur yang tersedia bagi pengguna terus mengalami perubahan.

Salah satu platform terbaru yang mendapatkan perhatian besar di kalangan pengguna muda adalah *Threads*, yang diluncurkan oleh *Instagram*. Threads memungkinkan pengguna untuk membuat postingan hingga 500 karakter dan dilengkapi dengan berbagai fitur yang mirip dengan X. Secara umum, Threads berfungsi serupa dengan X, yakni sebagai wadah bagi pengguna untuk berbagi konten berbasis teks. Meskipun ada kesamaan, tampilan dan fitur *Threads* sedikit berbeda dari X. Beberapa fitur yang ditawarkan Threads termasuk melihat profil, memeriksa mention, dan lain-lain. Berdasarkan data Google Play Store (2026), Threads memperoleh rating 4,5 dengan lebih dari 500 juta unduhan, sedangkan aplikasi X memperoleh rating yang lebih rendah yaitu 3,8 dengan lebih dari 1 miliar unduhan. Kondisi ini menunjukkan adanya perbedaan persepsi pengguna terhadap kedua platform tersebut. Perbedaan kualitas pengalaman pengguna tercermin pada rating yang diberikan melalui *Google Play Store*. *Rating* bukan sekadar angka statistik, melainkan bentuk nyata dari persepsi dan tingkat kepuasan pelanggan terhadap suatu aplikasi (A. D. Fitriani et al., 2026). Aplikasi dengan rating yang lebih tinggi umumnya menunjukkan bahwa mayoritas pengguna memberikan penilaian positif terhadap layanan yang diterima sehingga berpotensi meningkatkan kepercayaan calon pengguna untuk mengunduh aplikasi tersebut.

Penelitian yang dilakukan oleh Lim & Ayunda (2023) menjelaskan bahwa kemunculan Threads sebagai aplikasi baru dari Meta menarik perhatian masyarakat karena menawarkan konsep media sosial berbasis percakapan yang banyak dibandingkan dengan X. Kehadiran aplikasi ini memunculkan beragam tanggapan dari pengguna media sosial sehingga *Threads* menjadi salah satu objek penelitian yang relevan dalam kajian analisis sentimen maupun evaluasi pengalaman pengguna.

Analisis sentimen merupakan salah satu cabang dari *Natural Language Processing* (NLP) yang bertujuan untuk mengidentifikasi, mengekstraksi, dan mengklasifikasikan opini atau ekspresi emosi yang terkandung dalam data berbentuk teks (Hakim et al., 2025). Analisis sentimen penelitian ini berbasis machine learning. Pendekatan ini bekerja dengan cara melatih sebuah model menggunakan data latih yang sudah diberi label sentimen, sehingga model mampu mengenali pola dan karakteristik teks berdasarkan kategori tertentu. Setelah proses pelatihan selesai, model tersebut dapat digunakan untuk memprediksi sentimen dari data baru yang belum diberi label.

Penelitian yang dilakukan oleh (Amelia & Yustiana, 2024) dalam jurnal "Analisis Sentimen Pada Ulasan Produk Uniqlo dengan Algoritma *Naive Bayes*". penelitian ini, menggunakan metode analisis sentimen dengan algoritma *Naive Bayes* telah memberikan pemahaman yang lebih mendalam terhadap sentimen pelanggan terhadap produk *Uniqlo*. Profil responden menunjukkan mayoritas adalah perempuan dengan dominasi usia 23 tahun,

dan meskipun pengalaman situasional terjadi sekali, tingkat kenyamanan dan kepuasan pelanggan terkait situasi tersebut cukup tinggi. Analisis sentimen produk *Uniqlo* menggunakan *Naive Bayes* memberikan wawasan tentang sejauh mana sentimen positif, netral, dan negatif yang diberikan oleh pelanggan.

Penelitian yang dilakukan oleh Huwaida et al. (2024) dalam jurnal "Analisis Sentimen Komentar *Youtube* Terhadap Pemindahan Ibu Kota Negara Menggunakan Metode *Naive Bayes*". Hasil penerapan metode *Naive Bayes* dengan teknik SMOTE dalam mengklasifikasikan data komentar YouTube terhadap Ibu Kota Negara Nusantara menjadi kelas negatif, netral, dan positif dengan menggunakan perbandingan *data train* dan *data test* sebesar 90%:10% diperoleh hasil klasifikasi sentimen sebanyak 266 komentar diklasifikasikan dengan benar yang terdiri dari 17 komentar negatif, 213 komentar netral, dan 29 komentar positif dari total data sebanyak 324 komentar pada data test.

Penelitian yang dilakukan oleh Nurian et al. (2024) dalam jurnal "Analisis Sentimen Pengguna Aplikasi *Shopee* Pada Situs *Google Play* Menggunakan *Naive Bayes Classifier*". Hasil *accuracy* tertinggi adalah klasifikasi menggunakan *Naive Bayes Classifier* dengan seleksi fitur TF-IDF yaitu sebesar 77%, precision 26%, recall 33%, dan f1-*score* 29%, data testing yang digunakan sebanyak 300 data atau 10% dari 3000 data, dari jumlah data yang digunakan dengan metode acak pada saat testing.

Penelitian yang dilakukan oleh Rifa'I et al. (2024) dalam jurnal "Analisis Sentimen Terhadap Layanan Aplikasi *Grab* Indonesia Menggunakan Metode *Naive Bayes*". Dari hasil penelitian terhadap analisis sentimen pengguna aplikasi *Grab* menggunakan 2000 data teks yang di-crawling didapatkan hasil klasifikasi positif sebanyak 1677 dan sentimen negatif sebanyak 69 data teks artinya pengguna banyak memberikan komentar positif terhadap aplikasi *Grab*. Dan hasil klasifikasi menggunakan metode *Naive Bayes* didapatkan hasil *accuracy* sebesar 84,36% *precision* 99,64% dan *recall* 87,12%.

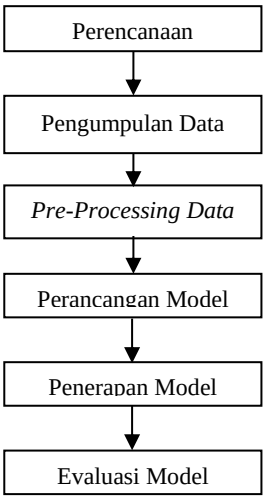
Penelitian yang dilakukan oleh Riskawati et al. (2024) dalam jurnal "Penerapan Metode *Naive Bayes Classifier* pada Analisis Sentimen Aplikasi *Gopay*. Dari data penelitian yang diambil melalui proses web scraping terhadap 1000 ulasan aplikasi *GoPay* dari 16 Agustus sampai 26 November 2023 dengan menggunakan perbandingan *data train* dan *data test* sebesar 70%:30%. Dari segi performa, hasil perhitungan menunjukkan tingkat akurasi sebesar 85%. Selain itu, terdapat nilai *precision* sebesar 84%, *recall* sebesar 96% dan f1-*score* sebesar 90%.

Tujuan penelitian ini membantu mengelompokkan ulasan berdasarkan polaritas yang terkandung di dalamnya sehingga tanggapan pengguna dapat dipahami secara lebih terstruktur yang dikelompokkan menjadi tiga kelas sentimen, yaitu positif, negatif, dan netral. Algoritma *Naive Bayes* digunakan untuk melakukan analisis sentimen terhadap ulasan pengguna aplikasi *Threads* yang diperoleh dari *Google Play Store*. Penelitian ini diharapkan dapat memberikan gambaran mengenai kecenderungan sentimen pengguna terhadap aplikasi *Threads* sekaligus mengetahui kemampuan algoritma *Naive Bayes* dalam melakukan analisis sentimen pada ulasan pengguna.

METODE

Penelitian ini menggunakan pendekatan kuantitatif. Pendekatan tersebut digunakan karena penelitian mengolah data ulasan pengguna aplikasi Threads dalam bentuk data yang dapat dihitung dan diukur. Data berupa teks ulasan pengguna aplikasi yang diproses melalui beberapa tahapan untuk memperoleh informasi mengenai sentimen yang ada di dalamnya. Hasil pengolahan kemudian digunakan untuk membentuk model menggunakan algoritma

Naive Bayes dan mengukur kinerjanya berdasarkan hasil prediksi. Berikut tahapan penelitian yang dituangkan pada gambar 1:



Gambar 1. Tahapan Penelitian
Sumber : Dokumentasi Pribadi

Perencanaan

Pada tahap ini peneliti menentukan permasalahan yang akan dikaji, tujuan penelitian, objek penelitian, sumber data, metode pengolahan data, algoritma yang digunakan, serta metode evaluasi. Objek yang digunakan adalah ulasan pengguna aplikasi *Threads* pada *Google Play Store*, sedangkan algoritma *Naive Bayes* digunakan sebagai metode untuk menentukan kelas sentimen pada ulasan pengguna berdasarkan fitur teks yang telah divektorisasi. Proses pengolahan dan penerapan metode dilakukan menggunakan komputasi yang mendukung pengolahan data serta pemodelan berbasis *Python*, yaitu *Google Colab*. Tahap ini menjadi dasar dalam menentukan rangkaian proses yang diterapkan pada data penelitian.

Pengumpulan data

Pada tahap ini, yang dikumpulkan yaitu ulasan pengguna aplikasi Threads yang tersedia di *Google Play Store*. Proses pengumpulan dilakukan dengan memanfaatkan teknik *web scraping* menggunakan *Google Colab* dan bahasa pemrograman Python serta pustaka pendukung. Sebanyak 5.000 data ulasan pengguna terbaru dari tanggal 9 November 2025 sampai tanggal 16 Agustus 2026 berhasil dikumpulkan sebagai data awal.

Pre-processing data

Pada tahap ini dilakukan pengolahan terhadap data ulasan agar menjadi data yang lebih terstruktur dan siap digunakan dalam proses analisis. Tahapan yang dilakukan meliputi *cleaning*, *case folding*, normalisasi, tokenisasi, *stopword removal*, dan *stemming*. Setelah itu, ulasan diberi label sentimen berdasarkan *rating*, kemudian dilakukan vektorisasi TF-IDF untuk mengubah teks menjadi fitur numerik yang dapat digunakan dalam penerapan model *Naive Bayes*.

Perancangan model

Pada tahap perancangan model ditentukan algoritma *Naive Bayes* yang digunakan untuk mengelompokkan ulasan ke dalam tiga kelas sentimen, yaitu positif, negatif, dan netral. Model dirancang menggunakan hasil vektorisasi TF-IDF sebagai fitur masukan, dengan pembagian data menjadi data latih dan data uji menggunakan rasio 80:20. Perancangan ini menjadi dasar dalam proses pelatihan dan pengujian model. Secara garis besar Teorema Bayes memiliki bentuk umum sebagai berikut:

$$P(C|X) = \frac{P(X|C) \times P(C)}{P(X)} \quad (1)$$

Penerapan model

Pada tahap penerapan model algoritma *Naive Bayes* diterapkan pada data yang telah melalui preprocessing dan vektorisasi TF-IDF. Data latih digunakan untuk membangun model berdasarkan pola kata pada setiap kelas sentimen, kemudian model digunakan untuk memprediksi kelas sentimen pada data uji. Hasil prediksi tersebut selanjutnya digunakan pada tahap evaluasi untuk mengetahui kemampuan model dalam mengelompokkan ulasan.

Evaluasi model

Pada tahap evaluasi model dilakukan pengukuran terhadap hasil prediksi algoritma *Naive Bayes* untuk mengetahui

kemampuan model dalam menentukan sentimen ulasan. Evaluasi dilakukan menggunakan *confusion matrix* dengan membandingkan hasil prediksi model terhadap label sentimen sebenarnya pada data uji. confusion matrix terdiri atas empat komponen, yaitu *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN). *True Positive* (TP). Menurut Cabot & Ross (2023), *confusion matrix* menjadi dasar dalam menghitung berbagai metrik evaluasi klasifikasi karena mampu menunjukkan hubungan antara hasil prediksi model dan kondisi aktual setiap kelas. Beberapa metrik evaluasi yang umum digunakan, yaitu *accuracy*, *precision*, *recall*, dan F1-*score*. Hasil evaluasi digunakan untuk mengetahui tingkat kinerja model dalam menentukan sentimen positif, negatif, dan netral.

a. *Accuracy* digunakan untuk mengukur tingkat ketepatan model dalam mengklasifikasikan seluruh data.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

b. *Precision* digunakan untuk mengukur ketepatan model ketika memberikan prediksi pada kelas positif.

$$Precision = \frac{TP}{TP+FP} \quad (3)$$

c. *Recall* digunakan untuk mengukur kemampuan model dalam mengenali seluruh data yang benar-benar termasuk kelas positif.

$$Recall = \frac{TP}{TP+FN} \quad (4)$$

d. F1-*score* merupakan ukuran yang menggabungkan nilai *precision* dan *recall* sehingga menghasilkan satu nilai evaluasi yang seimbang.

$$F1\text{-}Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (5)$$

HASIL DAN PEMBAHASAN

Analisis Data

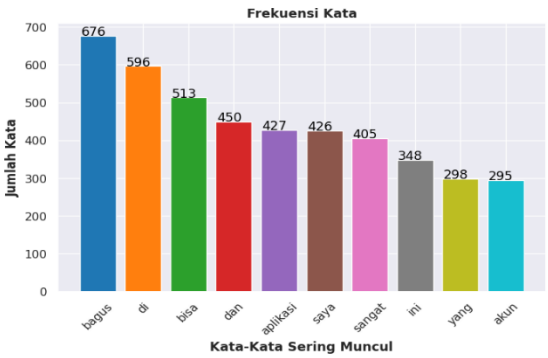
Hasil pengumpulan data ulasan dari *Google Play Store*, diperoleh sebanyak 5.000 data ulasan pengguna aplikasi Threads sebagai data awal. Data yang dikumpulkan terdiri atas beberapa atribut, yaitu ID pengguna, nama pengguna, rating, teks ulasan, dan tanggal. Setiap atribut memiliki jumlah data yang sama, yaitu 5.000 data, sehingga seluruh informasi tersebut dapat digunakan sebagai dataset awal dalam penelitian.

Tabel 1. Representasi Data

Rating	Ulasan
5	mantap
5	OK
5	mohon dg sangat berikan layanan terbaiknya dlm
5	Keren
5	sangat baguss
1	isinya robot semua, sekarang dikit-dikit akun di
1	tidak bisa akses
4	tambahkan fitur: terjemahkan otomatis komentar berbahasa
1	mohon untuk pengembangan developer aplikasi, kembalikan fitur aksesibilitas sehingga dapat dibaca oleh pengguna screen reader,
3	kasih bintang 3 dulu. kl Bagus d tambah

Gambar 2. *Word Cloud* Data Ulasan Pengguna

Sumber : Dokumentasi Pribadi



Gambar 3. Frekuensi Kemunculan 10 Kata Terbanyak

Sumber : Dokumentasi Pribadi

Pre-Processing Data

Text Pre-Processing

a. Cleaning

Tabel 2. *Cleaning*

Sebelum	Sesudah
OK	OK
isinya robot semua, sekarang dikit-dikit akun di tangguhkan	isinya robot semua sekarang dikit dikit akun di tangguhkan
tidak bisa akses	tidak bisa akses
tambahkan fitur: terjemahkan otomatis komentar berbahasa asing	tambahkan fitur terjemahkan otomatis komentar berbahasa asing
kasih bintang 3 dulu. kl Bagus d tambah	kasih bintang dulu. kl Bagus d tambah

b. Case folding

Tabel 3. *Case Folding*

Sebelum	Sesudah
OK	ok
isinya robot semua sekarang dikit dikit akun di tangguhkan	isinya robot semua sekarang dikit dikit akun di tangguhkan
tidak bisa akses	tidak bisa akses
tambahkan fitur terjemahkan otomatis komentar berbahasa asing	tambahkan fitur terjemahkan otomatis komentar berbahasa asing
kasih bintang dulu. kl Bagus d tambah	kasih bintang dulu. kl bagus d tambah

c. Normalisasi

Tabel 4. Normalisasi

Sebelum	Sesudah
ok	oke
isinya robot semua sekarang dikit dikit akun di tangguhkan	isinya robot semua sekarang sedikit sedikit akun di tangguhkan
tidak bisa akses	tidak bisa akses
tambahkan fitur terjemahkan otomatis komentar berbahasa asing	tambahkan fitur terjemahkan otomatis komentar berbahasa asing
kasih bintang dulu. kl bagus d tambah	kasih bintang dulu kalau bagus di tambah

d. Tokenisasi

Tabel 5. Tokenisasi

Sebelum	Sesudah
oke	['oke']
isinya robot semua sekarang sedikit sedikit akun di tangguhkan	['isinya', 'robot', 'semua', 'sekarang', 'sedikit', 'sedikit', 'akun', 'di',
tidak bisa akses	['tidak', 'bisa',
tambahkan fitur terjemahkan otomatis komentar berbahasa asing	['tambahkan', 'fitur', 'terjemahkan', 'otomatis',
kasih bintang dulu kalau bagus di tambah	['kasih', 'bintang', 'dulu', 'kalau', 'bagus', 'di', 'tambah']

e. Stopword removal

Tabel 6. *Stopword Removal*

Sebelum	Sesudah
['oke']	['oke']
['isinya', 'robot', 'semua', 'sekarang', 'sedikit', 'sedikit', 'akun', 'di', 'tangguhkan']	['isinya', 'robot', 'akun', 'tangguhkan']
['tidak', 'bisa',	['akses']
['tambahkan', 'fitur', 'terjemahkan', 'otomatis',	['tambahkan', 'fitur', 'terjemahkan', 'otomatis',
['kasih', 'bintang', 'dulu', 'kalau', 'bagus', 'di', 'tambah']	['kasih', 'bintang', 'bagus']

f. Stemming

Tabel 7. *Stemming*

Sebelum	Sesudah
['oke']	oke
['isinya', 'robot', 'semua', 'sekarang', 'sedikit', 'sedikit', 'akun', 'di', 'tangguhkan']	isi robot akun tangguh
['tidak', 'bisa',	akses
['tambahkan', 'fitur', 'terjemahkan', 'otomatis', 'komentar', 'berbahasa', 'asing']	tambah fitur terjemah otomatis komentar bahasa asing
['kasih', 'bintang', 'dulu', 'kalau', 'bagus', 'di', 'tambah']	bagus

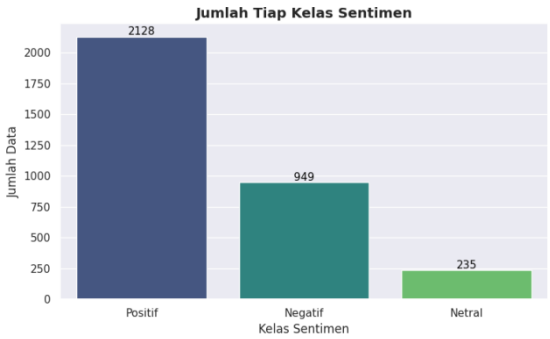
Setelah seluruh tahapan text preprocessing selesai dilakukan, diperoleh data ulasan yang telah melalui proses pembersihan dan penyesuaian format teks sesuai dengan kebutuhan penelitian. Data hasil pengolahan tersebut selanjutnya diperiksa untuk memastikan tidak terdapat nilai kosong yang dapat mengganggu proses analisis. Jumlah data yang diperoleh setelah *text preprocessing* menjadi 3312 data ulasan pengguna.

Pelabelan Sentimen Berdasarkan *Rating* Ulasan

Rating digunakan sebagai dasar untuk menentukan kelas sentimen karena setiap ulasan yang dikumpulkan memiliki nilai penilaian dalam rentang satu sampai lima bintang. Pendekatan ini memungkinkan setiap ulasan memperoleh label sentimen secara konsisten sebelum digunakan dalam proses klasifikasi. Sebanyak 3.312 data ulasan yang telah melewati proses text preprocessing diberi label sentimen berdasarkan rating yang diberikan pengguna pada *Google Play Store*. Dalam penelitian ini, rating dikelompokkan menjadi tiga kelas sentimen, yaitu positif, netral, dan negatif. Ulasan dengan *rating* 1 dan 2 bintang diberikan label negatif, *rating* 3 bintang diberikan label netral, sedangkan *rating* 4 dan 5 bintang diberikan label positif.

Tabel 8. Pelabelan Sentimen

Rating	Ulasan	Sentimen
5	oke	Positif
1	isi robot akun	Negatif
1	akses	Negatif
4	tambah fitur terjemah otomatis komentar bahasa	Positif
3	kasih bintang	Netral



Gambar 3. Jumlah Tiap Kelas Sentimen

Berdasarkan hasil pelabelan sentimen menggunakan rating, dari 3.312 data ulasan yang telah melalui tahap text preprocessing, terdapat 2.128 ulasan positif (64,25%), 949 ulasan negatif (28,65%), dan 235 ulasan netral (7,09%). Hasil tersebut menunjukkan bahwa data penelitian didominasi oleh sentimen positif, sedangkan sentimen netral memiliki jumlah paling sedikit. Distribusi tersebut menunjukkan adanya ketidakseimbangan kelas sebelum data digunakan dalam proses pemodelan. Kondisi ini perlu diperhatikan karena jumlah data yang tidak seimbang dapat menyebabkan model lebih banyak mempelajari karakteristik kelas mayoritas dibandingkan kelas dengan jumlah data yang lebih sedikit.

Vektorisasi TF-IDF

Vektorisasi TF-IDF terhadap teks ulasan untuk memberikan nilai pada setiap kata berdasarkan frekuensi kemunculannya dalam dokumen dan penyebarannya pada

keseluruhan dokumen. Kemudian diperoleh hasil perhitungan TF-IDF sebagai berikut:

Tabel 9. Perhitungan TF-IDF

Kata	TF-IDF
bagus	249.198073
aplikasi	186.718058
ya	105.602357
Threads	88.465557
mantap	86.372955
suka	75.881332
banget	71.962899
oke	71.114124
akun	69.924675
buka	65.639953
keren	57.233996
coba	45.567618
masuk	44.117382
baik	43.567635
teman	41.739567
bantu	41.372168
orang	41.210483
instagram	37.324827
tolong	36.872894
manfaat	36.732380

Pembagian data latih dan data uji

Setelah proses vektorisasi menggunakan TF-IDF, data dibagi menjadi data latih dan data uji dengan perbandingan 80:20. Pembagian ini dilakukan untuk menyediakan data yang digunakan dalam proses pembelajaran model sekaligus data yang digunakan untuk menguji kemampuan model dalam memberikan prediksi terhadap data yang belum digunakan selama pelatihan. Dari 3.312 data ulasan, diperoleh data latih dan data uji sebagai berikut:

Tabel 10. Pembagian Data Uji dan Data Latih

Data	Positif	Negatif	Netral	Total
Data latih	1702	759	188	2649
Data uji	426	190	47	663
Total	2128	849	235	3312

Pembagian tersebut menunjukkan bahwa distribusi kelas pada data latih belum seimbang, terutama pada kelas netral yang hanya berjumlah 188 data. Kondisi ini berpotensi menyebabkan model lebih banyak mempelajari pola dari kelas positif sebagai kelas dengan jumlah data terbesar, sementara karakteristik kelas netral lebih sedikit direpresentasikan dalam proses pembelajaran. Pada tahap berikutnya, *random oversampling* dilakukan terhadap data latih sehingga jumlah masing-masing kelas menjadi 1.702 data, terdiri atas 1.702 positif, 1.702 negatif, dan 1.702 netral. *Random Oversampling* merupakan penambahan data dari kelas minoritas ke dalam data training secara acak (R. D. Fitriani et al., 2021). Proses penambahan ini diulang sampai jumlah data kelas minoritas sama dengan jumlah kelas mayoritas. Dengan demikian, jumlah data latih setelah *oversampling* menjadi 5.106 data, sedangkan data uji berjumlah 663 tetap dipertahankan dalam kondisi awal agar evaluasi dapat menggambarkan kemampuan model terhadap distribusi data yang sebenarnya. Data latih hasil oversampling kemudian digunakan untuk melatih algoritma *Naive Bayes*, sementara data uji digunakan untuk menghasilkan prediksi dan mengevaluasi kinerja model melalui *confusion matrix*.

Penerapan Algoritma *Naive Bayes*

Algoritma *Naive Bayes* bertujuan untuk memprediksi probabilitas kejadian di masa depan berdasarkan pengalaman di masa lalu, sehingga sering disebut sebagai penerapan Teorema Bayes dalam konteks klasifikasi. Teorema tersebut dikombinasikan dengan *Naive* dimana diasumsikan kondisi antar atribut saling bebas (Ramdan et al., 2024). Sebelum proses pelatihan, ditemukan ketidakseimbangan jumlah data pada data latih. Kondisi tersebut menyebabkan model pada pengujian awal tidak mampu menghasilkan prediksi untuk kelas netral. Untuk mengatasi kondisi tersebut, dilakukan oversampling pada data. Data latih hasil oversampling selanjutnya digunakan untuk membentuk model *Naive Bayes*, sementara data uji digunakan untuk

mengetahui kemampuan model dalam memprediksi kelas sentimen.

Pada proses penerapannya, *Naive Bayes* menghitung probabilitas setiap kelas berdasarkan data latih dan vektorisasi TF-IDF dari setiap fitur. Model kemudian membandingkan probabilitas kelas positif, negatif, dan netral pada setiap ulasan dan menetapkan kelas dengan probabilitas terbesar sebagai hasil prediksi. Dengan adanya penyeimbangan pada data latih, model memperoleh representasi yang lebih seimbang dari ketiga kelas sehingga kelas netral yang sebelumnya tidak muncul dalam prediksi mulai dapat dikenali. Hasil prediksi terhadap data uji selanjutnya digunakan pada tahap evaluasi untuk mengukur kemampuan model melalui *confusion matrix*, *precision*, *recall*, dan F1-*score*.

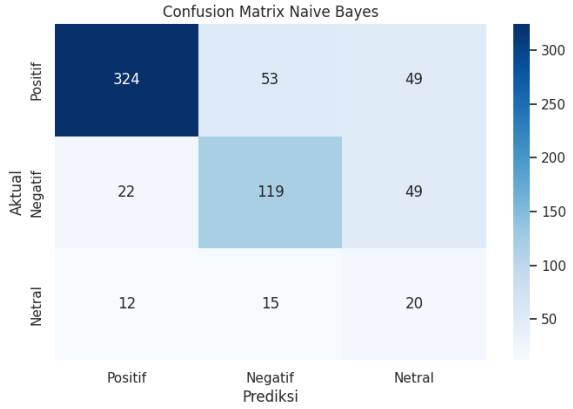
```
Label sebenarnya:
Sentimen
Positif    426
Negatif    190
Netral      47
Name: count, dtype: int64

Hasil prediksi:
Positif    358
Negatif    187
Netral     118
Name: count, dtype: int64
```

Gambar 4. Hasil Prediksi

Penerapan tersebut menghasilkan 663 prediksi pada data uji, dengan distribusi 358 ulasan diprediksi sebagai positif, 187 sebagai negatif, dan 118 sebagai netral. Hasil ini menunjukkan bahwa model tidak lagi mengabaikan kelas netral setelah dilakukan *oversampling*. Namun, distribusi hasil prediksi belum menunjukkan bahwa seluruh kelas telah dikenali dengan tingkat ketepatan yang sama. Oleh karena itu, penilaian terhadap kinerja *Naive Bayes* tidak cukup hanya berdasarkan jumlah prediksi, tetapi perlu dilanjutkan dengan evaluasi menggunakan *confusion matrix* dan metrik kinerja pada masing-masing kelas.

Evaluasi Menggunakan Confusion Matrix

Untuk mengetahui kemampuan algoritma *Naive Bayes* dalam memprediksi tiga kelas sentimen, yaitu positif, negatif, dan netral, evaluasi model dilakukan menggunakan *confusion matrix*. *Confusion matrix* membandingkan label aktual pada data uji dengan label hasil prediksi model sehingga dapat diketahui jumlah data yang diprediksi dengan benar maupun yang masih mengalami kesalahan. Evaluasi ini digunakan untuk melihat kinerja model secara keseluruhan melalui *accuracy* serta kinerja pada setiap kelas melalui *precision*, *recall*, dan F1-*score*.



Gambar 5. Confusion Matrix

	precision	recall	f1-score	support
Negatif	0.64	0.63	0.63	190
Netral	0.17	0.43	0.24	47
Positif	0.91	0.76	0.83	426
accuracy			0.70	663
macro avg	0.57	0.60	0.57	663
weighted avg	0.78	0.70	0.73	663

Gambar 6. Hasil Evaluasi

Hasil evaluasi menunjukkan bahwa model memperoleh *accuracy* sebesar 70%. Pada kelas positif, model memperoleh *precision* 0,91, *recall* 0,76, dan F1-*score* 0,83, yang menunjukkan bahwa model memiliki kemampuan paling baik dalam mengenali sentimen positif. Pada kelas negatif, diperoleh *precision* 0,64, *recall* 0,63, dan F1-*score* 0,63, sehingga kinerja model dalam mengenali sentimen negatif berada pada tingkat yang lebih rendah dibandingkan kelas positif. Sementara itu, kelas netral memperoleh *precision* 0,17, *recall* 0,43, dan F1-*score* 0,24, yang

menunjukkan bahwa kelas ini masih menjadi bagian yang paling sulit dikenali oleh model.

Nilai tersebut memperlihatkan bahwa penerapan random oversampling pada data latih membantu model untuk mulai mengenali kelas netral, tetapi kemampuan prediksinya masih belum optimal. *Recall* sebesar 0,43 menunjukkan bahwa model berhasil mengenali sebagian data yang sebenarnya termasuk sentimen netral, sedangkan *precision* sebesar 0,17 menunjukkan bahwa sebagian besar data yang diprediksi sebagai netral ternyata berasal dari kelas lain. Dengan demikian, meskipun ketidakseimbangan kelas telah ditangani pada data latih, karakteristik sentimen netral masih sulit dibedakan oleh model *Naive Bayes*. Hasil evaluasi ini selanjutnya menjadi dasar untuk menilai kinerja model secara keseluruhan dan mengidentifikasi kelas yang masih memerlukan perhatian dalam analisis sentimen ulasan *Threads*.

SIMPULAN

Berdasarkan hasil evaluasi menggunakan *confusion matrix* menunjukkan bahwa model memperoleh accuracy sebesar 70%. Kinerja terbaik diperoleh pada sentimen positif sedangkan sentimen netral memperoleh hasil yang lebih rendah. Hasil tersebut menunjukkan bahwa model *Naive Bayes* memiliki kemampuan yang lebih baik dalam mengenali sentimen positif dibandingkan dua kelas lainnya. Meskipun random oversampling membuat kelas netral mulai dapat diprediksi, model masih cenderung memberikan terlalu banyak prediksi pada kelas tersebut sehingga tingkat ketepatannya masih rendah.

Secara keseluruhan, algoritma *Naive Bayes* dapat diterapkan untuk menganalisis sentimen ulasan pengguna aplikasi Threads, dengan kinerja yang masih dipengaruhi oleh karakteristik dan distribusi data pada masing-masing kelas. Hasil penelitian menunjukkan bahwa penanganan ketidakseimbangan data membantu model mengenali kelas netral yang sebelumnya tidak muncul dalam prediksi, tetapi belum mampu menghasilkan kinerja yang seimbang pada seluruh kelas sentimen. Temuan tersebut dapat menjadi dasar untuk pengembangan penelitian selanjutnya dengan mempertimbangkan metode penyeimbangan data atau algoritma lain yang lebih mampu menangani karakteristik kelas netral.

DAFTAR PUSTAKA

Amelia, E. E., & Yustiana, I. (2024). Analisis Sentimen Pada Ulasan Produk Uniqlo dengan Algoritma Naive Bayes. *Jurnal Sains Komputer & Informatika (J-SAKTI*, *8*(1), 141–148.

Asosiasi Penyelenggara Jasa Internet Indonesia. (2026). *Survei Internet APJII 2026*. Asosiasi Penyelenggara Jasa Internet Indonesia (APJII). https://survei.apjii.or.id/

Badan Pusat Statistik. (2025). *Statistik Telekomunikasi Indonesia 2024*. Badan Pusat Statistik. https://www.bps.go.id/id/publication/2025/08/29/beaa2be400eda6ce6c636ef8/statistik-telekomunikasi-indonesia-2024.html

Cabot, J. H., & Ross, E. G. (2023). Evaluating prediction model performance. *Surgery (United States)*, *174*(3), 723–726. https://doi.org/10.1016/j.surg.2023.05.023

Fitriani, A. D., Sumawidjaja, R. N., Herlinawati, E., & Aziz, D. A. (2026). Pengaruh Kualitas Produk, Online Customer Review dan Online Customer Rating Terhadap Kepuasan Nasabah Dalam Menggunakan Aplikasi M Banking KB Star. *ECo-Buss*, *8*(3), 2256–2271. https://doi.org/10.32877/eb.v8i3.3478

Fitriani, R. D., Yasin, H., & Tarno. (2021). Penanganan Klasifikasi Kelas Data Tidak Seimbang dengan Random

Oversampling pada Naive Bayes. *Gaussian*, *10*, 11–20. https://ejournal3.undip.ac.id/index.php/gaussian/

Google Play Store. (2026). *Threads*. Google Play. https://play.google.com/store/apps/details?id=com.instagram.barcelona&hl=id

Hakim, M. J., Arya, M., Nugroho, S., Delpiero, A., Dhiaulhaq, M. H., & Rizal, K. (2025). *Analisa Review Netflix di Google Play Menggunakan Naïve Bayes*. *7*(3), 461–472.

Huwaida, S. F., Kusumawati, R., & Isnaini, B. (2024). Analisis Sentimen Komentar YouTube terhadap Pemindahan Ibu Kota Negara Menggunakan Metode Naïve Bayes. *Jambura Journal of Informatics*, *6*(1), 26–39. https://doi.org/10.37905/jji.v6i1.24718

Lim, F., & Ayunda, A. T. (2023). Sentiment Analysis of Indonesian People Toward Threads Application on Twitter Using Naïve Baiyes. *JUSIFO (Jurnal Sistem Informasi)*, *9*(2), 55–64. https://jurnal.radenfatah.ac.id/index.php/jusifo/article/view/19965

Nurian, A., Ma'arif, M. S., Amalia, I. N., & Rozikin, C. (2024). Analisis Sentimen Pengguna Aplikasi Shopee Pada Situs Google Play Menggunakan Naive Bayes Classifier. *Jurnal Informatika Dan Teknik Elektro Terapan*, *12*(1). https://doi.org/10.23960/jitet.v12i1.3631

Ramdan, M., Surya, A., & Hayati, U. (2024). Analisis Sentimen Ulasan Pengguna Ovo Menggunakan. *JATI(Jurnal Mahasiswa Teknik Informatika)*, *8*(3), 2780–2786.

Ramdani, A., Satria, A., Putri, A., Alia, H., & Zahra, A. (2026). *Pemanfaatan Media Sosial Terhadap Partisipasi Kewargaan Digital Generasi Z Di Indonesia : Literature Review*. *3*(1), 1–9.

Rifa'I, A., Ardhani, R., Pratama, D., & Fatihanursari, F. (2024). Analisis Sentimen Terhadap Layanan Aplikasi Grab Indonesia Menggunakan Metode Naïve Bayes. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *8*(1), 303–309. https://doi.org/10.36040/jati.v8i1.8425

Riskawati, R., Fatihanursari, F., Iin, I., & Rizki Rinaldi, A. (2024). Penerapan Metode Naïve Bayes Classifier Pada Analisis Sentimen Aplikasi Gopay. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *8*(1), 346–353. https://doi.org/10.36040/jati.v8i1.8699

Yusup, A, H., Azizah, A., Reejeki, Endang, S., & Meliza, S. (2023). Literature Review: Peran Media Pembelajaran Berbasis Augmented Reality Dalam Media Sosial. *JPI: Jurnal Pendidikan Indonesia*, *2*(5), 1–13. https://doi.org/10.59818/jpi.v3i5.575