Jurnal Statistika Dan Komputasi (STATKOM), Vol. 4 No. 2 (Dec 2025)
ISSN : 2963-038X (print), ISSN : 2963-0398
DOI : https://doi.org/10.32665/statkom.v4i2.5194

Analisis Sentimen Ulasan Pengguna BRImo Terhadap Pembaruan Fitur Aplikasi Menggunakan Naive Bayes Dengan Seleksi Fitur Chi-Square

Heppy Nur Asavia Ginasputri[1], Atika Dwi Saputri[2], Siti Nurasriyanti Wahid[3], Ariska Fitriyana Ningrum[4], M. Al Haris[5]

[1,2,3,5]Program Studi Statistika, Universitas Muhammadiyah Semarang
[4]Program Studi Sains Data, Universitas Muhammadiyah Semarang
E-mail: heppynag23@gmail.com

Diajukan 21 Juli 2025 ***Diperbaiki*** 14 Desember 2025 ***Diterima*** 21 Desember 2025

Abstrak

Latar Belakang: Aplikasi BRImo merupakan layanan mobile banking dari Bank Rakyat Indonesia (BRI) dengan jumlah pengguna yang besar. Ulasan pengguna di *Google Play Store* menjadi sumber data penting untuk mengetahui persepsi dan tingkat kepuasan pengguna, namun jumlah dan keragaman teks ulasan memerlukan metode analisis sentimen yang tepat.
Tujuan: Penelitian ini bertujuan mengevaluasi kinerja algoritma Naive Bayes dalam klasifikasi sentimen ulasan BRImo serta pengaruh seleksi fitur Chi-Square.
Metode: Metode penelitian meliputi praproses data teks yang terdiri dari *cleaning*, *case folding*, *tokenizing*, *normalization*, *filtering*, dan *stemming*. Selanjutnya dilakukan pembobotan fitur menggunakan TF-IDF dan seleksi fitur dengan metode Chi-Square, kemudian klasifikasi sentimen menggunakan algoritma Naive Bayes. Evaluasi model dilakukan menggunakan *confusion matrix* dengan metrik akurasi, presisi, *recall*, dan *F1-score*.
Hasil: Hasil penelitian menunjukkan bahwa model klasifikasi sentimen mencapai akurasi sebesar 95%, presisi 98%, *recall* 96%, dan *F1-score* 97%. Nilai recall yang tinggi menunjukkan kemampuan model yang sangat baik dalam mendeteksi sentimen positif pada ulasan pengguna.
Kesimpulan: Algoritma Naive Bayes dengan seleksi fitur Chi-Square efektif dalam menganalisis sentimen ulasan aplikasi BRImo dan dapat digunakan sebagai dasar evaluasi pengembangan aplikasi, namun kinerjanya masih terbatas dalam mendeteksi sentimen negatif akibat ketidakseimbangan data.

Kata kunci: Analisis Sentimen, BRImo, Naive Bayes, Chi-Square.

Abstract

Background: *The BRImo app is a mobile banking service from Bank Rakyat Indonesia (BRI) with a large number of users. User reviews on Google Play Store are an important source of data for understanding user perceptions and satisfaction levels, but the number and diversity of review texts require an appropriate sentiment analysis method.*
Objective: *This study aims to evaluate the performance of the Naive Bayes algorithm in classifying BRImo reviews by sentiment and the effect of Chi-Square feature selection.*
Methods: *The research method includes text data preprocessing consisting of cleaning, case folding, tokenizing, normalization, filtering, and stemming. Next, feature weighting is performed using TF-IDF and feature selection using the Chi-Square method, followed by sentiment classification using the Naive Bayes algorithm. Model evaluation is performed using a confusion matrix with accuracy, precision, recall, and F1-score metrics.*
Results: *The results show that the sentiment classification model achieved an accuracy of 95%, precision of 98%, recall of 96%, and an F1-score of 97%. The high recall value indicates the model's excellent ability to detect positive sentiment in user reviews.*
Conclusion: *The Naive Bayes algorithm with Chi-Square feature selection is effective in analyzing BRImo application review sentiment and can be used as a basis for evaluating application development, but its performance is still limited in detecting negative sentiment due to the quantity of sentiment data.*

Keywords : *Sentiment Analysis, BRImo, Naive Bayes, Chi-Square.*

PENDAHULUAN

Perkembangan teknologi informasi yang semakin cepat mendorong industri perbankan untuk berinovasi dalam meningkatkan mutu layanan. Salah satu bentuk inovasi tersebut adalah mobile banking, yang memungkinkan nasabah melakukan berbagai transaksi secara mudah melalui perangkat seluler. Seiring meningkatnya kebutuhan masyarakat akan layanan keuangan yang cepat, aman, dan praktis, digitalisasi perbankan menjadi keharusan agar bank tetap kompetitif dan mampu memenuhi perubahan perilaku konsumen yang kini lebih mengutamakan layanan berbasis digital (Arminda et al., 2023).

Bank Rakyat Indonesia (BRI) turut mendorong transformasi digital melalui aplikasi BRImo, yang dikembangkan untuk menyediakan layanan keuangan yang praktis dan mudah diakses (Permana et al., 2022). Lebih dari 10 juta unduhan dan ribuan ulasan di *Google Play Store* menunjukkan tingginya penggunaan aplikasi ini, sekaligus menyediakan sumber data berharga untuk menilai kualitas dan kinerja layanannya (Insan et al., 2023).

Peningkatan jumlah pengguna BRImo menghasilkan banyak ulasan yang menggambarkan pengalaman, tingkat kepuasan, hingga keluhan pengguna. Informasi sentimen dalam ulasan tersebut dapat dianalisis untuk memahami persepsi publik terhadap layanan (Rosyadi et al., 2025). Oleh karena itu, diperlukan metode klasifikasi sentimen otomatis agar pengembang dapat memperoleh evaluasi yang lebih efisien dan mendukung perbaikan layanan secara tepat sasaran (Elysa et al., 2023).

Analisis sentimen digunakan untuk mengidentifikasi opini dalam teks dan mengklasifikasikan ulasan ke dalam sentimen positif, negatif, atau netral (Sidabutar, 2023). Dalam penelitian ini digunakan algoritma Naive Bayes karena memiliki keunggulan dalam kesederhanaan, efisiensi komputasi, serta performa yang baik pada data teks pendek seperti ulasan pengguna (Ratiasasadara et al., 2023). Naive Bayes bekerja berdasarkan pendekatan probabilistik dan meskipun mengasumsikan independensi antar fitur, metode ini terbukti mampu menghasilkan akurasi yang kompetitif dalam analisis sentimen (Habiba et al., 2023). Namun, data teks umumnya menghasilkan jumlah fitur yang sangat besar, sehingga diperlukan proses seleksi fitur seperti Chi-Square untuk memilih fitur yang paling relevan dan menjaga kinerja model tetap optimal.

Namun, pada data teks, jumlah fitur yang dihasilkan dari tokenisasi dan representasi seperti TF-IDF (*Term Frequency-Inverse Document Frequency*) sangat besar dan dapat mengandung *noise* atau fitur yang tidak relevan, sehingga mempengaruhi akurasi klasifikasi. Untuk itu, diperlukan proses seleksi fitur yang mampu menyaring fitur-fitur yang paling berpengaruh terhadap kelas target. Dalam penelitian ini digunakan metode Chi-Square, karena terbukti efektif dalam mengukur tingkat ketergantungan antara fitur dan label sentimen, serta meningkatkan performa model klasifikasi (Ernayanti et al., 2023).

Pemilihan kombinasi Naive Bayes dan Chi-Square juga didukung oleh berbagai studi sebelumnya. Penelitian oleh Wijaya (2021) menunjukkan bahwa metode Chi-Square berhasil meningkatkan akurasi model Naive Bayes dalam menganalisis sentimen masyarakat terhadap kebijakan publik. Selain itu, juga menekankan pentingnya pemilihan metode yang ringan dan *scalable* seperti Naive Bayes untuk data ulasan aplikasi yang terus bertambah secara dinamis.

Berbagai penelitian sebelumnya telah menerapkan Naive Bayes dan metode seleksi fitur dalam analisis sentimen, namun

mayoritas masih menggunakan *dataset* berskala kecil serta belum menguji kinerja kombinasi Naive Bayes dan Chi-Square pada *volume* data ulasan yang besar.

Dataset ulasan pengguna BRImo dari *Google Play Store* digunakan sebagai bahan analisis. berdasarkan penjelasaan latar belakang yang telah diuraikan sebelumnya, peneliti tertarik untuk melakukan penelitian yang berjudul "Klasifikasi Sentimen Menggunakan Naive Bayes Dengan Seleksi Fitur Chi-Square (Studi Kasus : Ulasan Pengguna Aplikasi BRImo)" bertujuan untuk mengkaji performa algoritma Naive Bayes dalam klasifikasi sentimen ulasan pengguna BRImo, dengan bantuan seleksi fitur Chi-Square, guna menghasilkan model klasifikasi yang efisien, akurat, dan mampu membantu pihak pengembang memahami persepsi publik terhadap aplikasi.

Penelitian oleh Chen et al. (2021) juga mendukung efektivitas algoritma Naive Bayes dalam klasifikasi data, dengan menunjukkan bahwa modifikasi seperti *feature weighting* dan *Laplace calibration* dapat meningkatkan akurasi hingga lebih dari 99% dan menjaga kestabilan meskipun pada *dataset* berskala besar. Hal ini memperkuat relevansi penerapan Naive Bayes dalam analisis ulasan aplikasi seperti BRImo, karena algoritma ini tidak hanya sederhana dan cepat, tetapi juga mampu menangani data teks dalam jumlah besar secara efisien dan akurat (Chen et al., 2021).

Berdasarkan kesenjangan penelitian tersebut, penelitian ini bertujuan mengkaji kinerja algoritma Naive Bayes dengan seleksi fitur Chi-Square dalam klasifikasi sentimen ulasan pengguna aplikasi BRImo, serta memberikan kontribusi berupa model klasifikasi yang efisien dan akurat.

METODE

Desain Penelitian

Penelitian ini merupakan penelitian kuantitatif menggunakan data sekunder. Penelitian ini mengkombinasikan metode klasifikasi Naive Bayes dengan seleksi fitur menggunakan Chi-Square. Pendekatan ini bertujuan untuk membangun model klasifikasi sentimen yang efisien dan akurat berdasarkan ulasan pengguna. Evaluasi model dilakukan dengan menggunakan metrik akurasi, presisi, *recall*, *F1-score*, dan *confusion matrix*. Metode ini memilih fitur (kata) yang memiliki tingkat ketergantungan signifikan terhadap kategori sentimen, sehingga mengurangi jumlah fitur tidak relevan dan meningkatkan akurasi model klasifikasi (Prastyo et al., 2021).

Populasi dan Sampel

Populasi dalam penelitian ini adalah seluruh ulasan pengguna aplikasi BRImo di *Google Play Store*. Pemilihan berdasarkan relevansi terhadap fokus penelitian, yakni ekspresi sentimen positif dan negatif yang telah melalui proses pembersihan dan praproses data. Sampel ini dianggap mewakili beragam opini, pengalaman, dan kepuasan pengguna aplikasi BRImo.

Teknik Sampling

Penelitian ini menggunakan teknik *purposive sampling* berdasarkan kriteria ulasan yang relevan, yaitu ulasan yang mengandung ekspresi sentimen (positif, negatif, atau netral). Ulasan dengan konten kosong, spam, atau tidak relevan dikeluarkan dari *dataset*.

Tabel 1. *Dataset* Penelitian

No	Score	Ulasan
1	5	Memudahkan untuk semua transaksi nasabah
2	5	Puas
3	5	lebih aman praktis simpel LG cepet dan mudah terima kasih aplikasi brimooo amanah selalu,,,
...	...	...
99.997	5	Sangat membantu
99.998	5	Mantap

Subjek Penelitian

Data yang digunakan dalam penelitian ini diambil dari *Google Play Store*, yang menyediakan ulasan dari pengguna aplikasi BRImo.

Teknik Analisis Data

Penelitian ini menggunakan metode Algoritma Naive Bayes yang dipadukan dengan seleksi fitur Chi-Square. Proses penelitian meliputi beberapa tahap, yaitu *Data Selection, Data Cleaning (cleaning, Case folding, tokenizing, Normalization, Filtering, Stemming), Data Transformation (TF–IDF Term Frequency–Inverse Document Frequency), Feature Selection (Chi-Square), Data Mining, Pattern Evaluation* (akurasi, presisi, *recall*, *F1-score, coffusion matrix),* dan *Knowledge Presentation* (Arminda et al., 2023). Berikut adalah ilustrasi mengenai rincian langkah-langkah analisis data yang akan diterapkan dalam penelitian ini:

1. *Data Selection*: Data diambil dari *Google Play Store*, dengan atribut yang digunakan adalah *score* dan ulasan. Validasi pelabelan dilakukan melalui pemeriksaan manual untuk mengurangi kesalahan klasifikasi otomatis (Andriani & Wibowo, 2021). Meskipun proses pelabelan dilakukan secara otomatis berdasarkan skor *rating*, pendekatan ini memiliki kelemahan karena tidak semua ulasan dengan skor tertentu sesuai dengan sentimen dalam teksnya. Kondisi ini dapat menyebabkan *label noise*, yaitu ketidaksesuaian antara label dan isi ulasan. Misalnya, beberapa pengguna memberikan skor tinggi tetapi menuliskan keluhan dalam teks ulasannya.
2. *Data Cleaning*: Praproses teks dilakukan untuk meningkatkan kualitas data ulasan melalui tahapan:
   a. *Cleaning*: Menghilangkan tanda baca, angka, dan karakter yang tidak relevan.
   b. *Case Folding*: Mengonversi seluruh teks menjadi huruf kecil.
   c. *Tokenizing*: Memecah teks menjadi unit kata (*token*).
   d. *Normalization*: Menstandarkan kata tidak baku atau slang menjadi kata baku.
   e. *Filtering*: Menghilangkan *stopwords* (kata-kata umum yang tidak memiliki makna).
   f. *Stemming*: Mengubah kata ke bentuk dasarnya. Tahapan ini mengacu pada prosedur standar praproses data teks untuk analisis sentimen (Saepudin et al., 2023).
3. *Data Transformation*: Teks yang telah dibersihkan diubah menjadi bentuk numerik dengan menggunakan metode TF-IDF (*Term Frequency-Inverse Document Frequency*). TF-IDF digunakan untuk memberikan bobot penting terhadap kata-kata yang lebih informatif dalam teks ulasan (Ernayanti et al., 2023).
4. *Feature Selection*: Pemilihan fitur dilakukan menggunakan metode Chi-Square. Metode ini memilih fitur (kata) yang menunjukkan tingkat ketergantungan yang signifikan terhadap kategori sentimen, sehingga dapat mengurangi jumlah fitur yang tidak relevan dan meningkatkan akurasi model klasifikasi (Wijaya, 2021). Gautam dan Waoo (2024) mengembangkan pendekatan *hibrida* dengan menggabungkan metode Chi-Square dan *Recursive Feature Elimination (RFE)* untuk seleksi fitur dalam analisis sentimen. Pendekatan ini menunjukkan peningkatan signifikan dalam akurasi, presisi, *recall*, dan skor F1 dibandingkan dengan metode seleksi fitur lainnya. (Gautam & Waoo, 2024)

   Metode Chi-Square digunakan untuk menghitung keterkaitan setiap fitur dengan kelas sentimen dan mengurutkannya berdasarkan nilai Chi-Square. Pada penelitian ini, seluruh 1001 fitur hasil ekstraksi TF-IDF tetap

digunakan tanpa pembatasan jumlah fitur, sehingga Chi-Square berfungsi sebagai pengurutan fitur. Pendekatan ini mengacu pada Ernayanti et al. (2023) yang menunjukkan bahwa Chi-Square efektif dalam mengidentifikasi fitur penting dalam klasifikasi teks.

5. *Data Mining*: Pada tahap ini, dilakukan klasifikasi data ulasan menggunakan algoritma Naive Bayes. Algoritma ini dipilih karena sederhana, cepat, dan efektif dalam menangani data teks pendek serta mampu memberikan akurasi tinggi dalam klasifikasi (Anisah et al., 2016).
6. Klasifikasi Naive Bayes **:** Model Naive Bayes pertama-tama dilatih tanpa seleksi fitur Chi-Square sebagai *baseline* untuk membandingkan kinerjanya dengan model akhir yang menggunakan Chi-Square, sehingga dapat diketahui pengaruh seleksi fitur terhadap performa klasifikasi. Data dibagi menggunakan *stratified sampling* dengan *random seed* 42 agar proporsi kelas pada data latih dan uji seimbang dan hasil penelitian dapat direproduksi. Validasi dilakukan melalui *train-test split* dengan proporsi 80% untuk pelatihan dan 20% untuk pengujian, di mana model dilatih menggunakan fitur TF-IDF yang pada model akhir telah diurutkan berdasarkan nilai Chi-Square.
7. *Pattern Evaluation*: *Confusion Matrix* adalah metode dalam *data mining* yang digunakan untuk menilai kinerja algoritma atau model klasifikasi. *Confusion* matrix menggambarkan hasil prediksi model terhadap data uji dalam format tabel, yang menunjukkan empat kemungkinan hasil klasifikasi (Farissa et al., 2021).

Tabel 2. *Confusion Matrix*

		Kelas Sebenarnya	
		True	False
Kelas Prediksi	True	TP	FP
	False	FN	TN

Keterangan:

a. *True Positive* (TP), model memprediksi sentimen positif dan label sebenarnya positif.
b. *True Negative* (TN), model memprediksi sentimen positif, tetapi label sebenarnya negatif.
c. *False Negative* (FP), model memprediksi sentimen negatif dan label sebenarnya negatif.
d. *False Negative* (FN), model memprediksi sentimen negatif, tetapi label sebenarnya positif.

Pengujian dengan metode *confusion matrix* menghasilkan empat jenis keluaran, yaitu:

a. Akurasi, merupakan proporsi klasifikasi yang benar terhadap total data. Berikut rumus menghitng akurasi:

$$Akurasi = \frac{TP+TN}{(TP+TN+FP+FN)} \quad (1)$$

b. Presisi, merupakan ketepatan prediksi terhadap masing-masing label sentimen. Berikut rumus menghitung presisi:

$$Presisi = \frac{TP}{TP+FP} \quad (2)$$

c. *Recall*, merupakan kemampuan model untuk mengidentifikasi semua data yang relevan. Berikut rumus menghitng recall:

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

d. *F1-Score*, merupakan rata-rata harmonis dari nilai precision dan recall. Evaluasi juga dilengkapi dengan pembuatan *confusion matrix* untuk analisis distribusi prediksi yang lebih rinci (Paembonan & Abduh, 2021). Berikut rumus menghitng *F1-Score*:

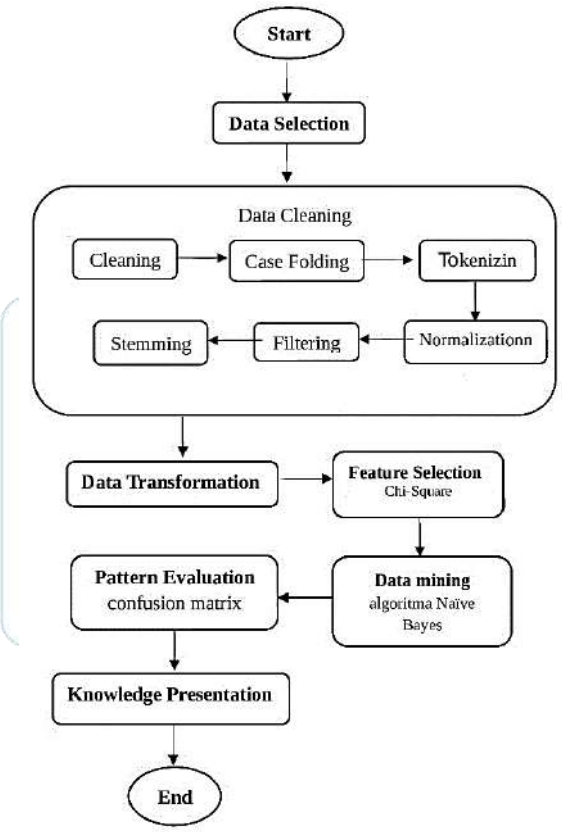
$$F1\ Score = 2\frac{(precision\ x\ recall)}{precison+recall} \quad (4)$$

Keterangan :

TP : Total prediksi *true positive*
TN : Total prediksi *true negative*
FP : Total prediksi *false positive*
FN : Total prediksi *false negative*

8. *Knowledge Presentation*: Hasil klasifikasi sentimen divisualisasikan menggunakan *wordcloud* untuk masing-masing kategori sentimen. *Wordcloud* digunakan untuk memperlihatkan kata-kata dominan dalam ulasan, sehingga dapat membantu memahami aspek utama yang menjadi perhatian pengguna (Farissa et al., 2021)

Adapun alur penelitian dapat dilihat pada Gambar 1.



Gambar 1. Metodologi Penelitian

HASIL DAN PEMBAHASAN

Deskripsi Data

Penelitian ini menggunakan dua variabel yang terdiri dari :

1. *Content* (isi ulasan) - Ulasan berupa teks yang diberikan oleh pengguna aplikasi BRImo.
2. *Score* (penilaian pengguna) - Nilai yang diberikan oleh pengguna dalam rentang skor 1 hingga 5.

Variabel-variabel ini digunakan untuk melakukan analisis sentimen terhadap ulasan pengguna aplikasi BRImo. Penentuan label sentimen dilakukan berdasarkan nilai skor dengan ketentuan berikut:

a. Skor 1-2: Sentimen negatif.
b. Skor 4-5: Sentimen positif.

Pemilihan variabel ini didasarkan pada relevansi data terhadap analisis sentimen, yang bertujuan untuk memahami kepuasan dan ketidakpuasan pengguna terhadap aplikasi BRImo.

Analisis Data

Penelitian ini dilakukan dengan menggunakan bahasa pemrograman *Python* di *platform Google Colaboratory* dan menerapkan algoritma *Multinomial Naive Bayes Classifier* untuk mengembangkan model analisis sentimen terhadap ulasan pengguna aplikasi BRImo yang diambil dari *Google Play Store*. Untuk meningkatkan akurasi klasifikasi, digunakan metode seleksi fitur Chi-Square guna menyaring fitur-fitur yang paling relevan. Proses pengembangan model ini mencakup tujuh tahapan sebagai berikut.

1. *Data Selection*:
   Dalam tahap ini, dilakukan pengambilan data ulasan pengguna aplikasi BRImo yang berbahasa Indonesia dari *Google Play Store*. Didapatkan 99.998 data yang akan di proses dalam penelitian ini.
2. Statistik Deskriptif:

Tabel 3. Statistik Deskriptif

Skor	Frekuensi	Persentase(%)
5	77543	77.544551
4	11228	11.228225
3	5944	5.944119
2	2813	2.813056
1	2469	2.469049

Berdasarkan Tabel 3, mayoritas pengguna BRImo memberikan skor 5 (sangat puas) dengan persentase sebesar 77,54%, menunjukkan bahwa sebagian besar pengguna merasa sangat puas terhadap layanan BRImo. Skor 4 (puas) juga memiliki persentase yang cukup signifikan yaitu

11,23%, yang berarti total 88,77% pengguna memiliki tingkat kepuasan tinggi (skor 4 dan 5). Sementara itu sebagian kecil pengguna memberikan skor rendah yaitu hanya 5,28% pengguna yang memberikan skor rendah yaitu 1 dan 2 menunjukkan bahwa ketidakpuasan terhadap layanan tergolong sangat rendah.

3. *Data Cleaning*:

Proses *data cleaning* dilakukan dengan tahapan-tahapan dalam prapemrosesan data yang meliputi:

a. *Cleaning*

Dilakukan penghapusan atribut yang tidak relevan dengan proses klasifikasi..

Tabel 4. *Cleaning*

Sebelum	Sesudah
Bagus tapi sayang buat beli pulsa dan kuota gabisa, tpi dah baguss lah	Bagus tapi sayang buat beli pulsa dan kuota gabisa tpi dah baguss lah

Dari Tabel 4, ditunjukkan proses cleaning (pembersihan teks) pada data ulasan pengguna. Pada kolom “Sebelum”, terdapat teks yang masih mengandung tanda baca seperti koma (,). Setelah proses cleaning, tanda baca tersebut dihapus, seperti yang terlihat di kolom “Sesudah”.

b. *Case Folding*

Dilakukan perubahan format karakter dokumen menjadi seluruhnya lowercase.

Tabel 5. *Case Folding*

Sebelum	Sesudah
Bagus tapi sayang buat beli pulsa dan kuota gabisa tpi dah baguss lah	bagus tapi sayang buat beli pulsa dan kuota gabisa tpi dah baguss lah

Berdasarkan Tabel 5, teks pada kolom "Sebelum" diubah menjadi huruf kecil semua pada kolom "Sesudah".

c. *Tokenizing*

Kalimat dipecah menjadi kata-kata (token) yang dipisahkan oleh spasi atau tanda koma.

.

Tabel 6. *Tokenizing*

Sebelum	Sesudah
bagus tapi sayang buat beli pulsa dan kuota gabisa tpi dah baguss lah	[bagus, tapi, sayang, buat, beli, pulsa, dan, kuota, gabisa, tpi, dah, baguss, lah]

Berdasarkan Tabel 6, terlihat proses pemecahan kalimat menjadi bagian-bagian yang lebih kecil (token).

d. *Normalization*

Dilakukan pengoreksian kata yang merupakan singkatan, tidak sesuai kaidah bahasa baku, atau mengandung kesalahan penulisan (*typo*).

Tabel 7. *Normalization*

Sebelum	Sesudah
[bagus, tapi, sayang, buat, beli, pulsa, dan, kuota, gabisa, tpi, dah, baguss, lah]	[bagus, tetapi, sayang, buat, beli, pulsa, dan, kuota, tidak bisa, tetapi, sudah, bagus, lah]

Berdasarkan Tabel 7, dilakukan proses normalisasi terhadap *tokens* tidak baku.

e. *Filtering*

Stopwords, berupa kata-kata yang kurang bermakna deskriptif, dihapus dari teks.

Tabel 8.*Filtering*

Sebelum	Sesudah
[bagus, tetapi, sayang, buat, beli, pulsa, dan, kuota, tidak bisa, tetapi, sudah, bagus, lah]	[bagus, pulsa, kuota, tidak bisa, bagus]

Berdasarkan Tabel 8, ditampilkan proses penyaringan kata-kata penting dari sebuah kalimat yang telah dilakukan normalisasi. Proses filtering atau *stopword removal* bertujuan untuk menghapus kata-kata umum (*stopwords*) yang tidak memiliki makna penting atau kontributif terhadap analisis, seperti: tetapi, dan, sudah, lah, buat, beli, dan lainnya.

f. *Stemming*

Kata-kata dalam dokumen ditransformasikan menjadi bentuk dasarnya

Tabel 9. *Stemming*

Sebelum	Sesudah
[bagus, pulsa, kuota, tidak bisa, bagus]	[bagus, pulsa, kuota, tidak bisa, bagus]

Hasil *stemming* pada Tabel 9. Menunjukkan bahwa tidak ada perubahan antara teks sebelum dan sesudah stemming. Hal ini terjadi karena semua kata seperti bagus, pulsa, kuota, dan frasa tidak bisa sudah merupakan bentuk dasar.

4. *Data Transformation*

Data yang telah melalui tahap pra-pemrosesan kemudian diberi label sebagai data positif dan negatif. Selanjutnya, baik dataset maupun label yang semula berbentuk teks diubah menjadi format numerik menggunakan algoritma TF–IDF sebagai berikut.

	abis	ad	ada	adalah	adanya	admin	agak	agar	aj	aja	...	wallet	\
0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	
1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	
2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	
3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	

	whatsapp	wifi	ya	yaa	yah	yang	yg	you	sentiment
0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	positif
1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	positif
2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	positif
3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	positif
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	positif

Gambar 2. Matriks hasil TF-IDF

Berdasarkan Gambar 2. Menunjukkan setiap kolom mewakili fitur kata (misalnya: "abis", "ada", "admin") dan nilai 0.0 menunjukkan bahwa kata tersebut tidak muncul dalam dokumen pada baris tersebut. Kolom terakhir "*sentiment*" berisi label sentimen untuk masing-masing ulasan.

5. *Feature Selection*

Pemilihan fitur dilakukan menggunakan metode Chi-Square, yang bertujuan untuk memilih kata-kata dengan tingkat keterkaitan signifikan terhadap kategori sentimen. Pendekatan ini membantu mengeliminasi fitur yang kurang relevan dan meningkatkan kinerja akurasi model klasifikasi. Fitur terbaik sebagai berikut.

Tabel 10. Fitur Terbaik

Fitur	*Chi-Square Score*
bisa	2421.253852
gagal	2081.866359
tidak	1885.833815
susah	1496.115824
mantap	1065.462926

Berdasarkan Tabel 10. Menunjukkan bahwa kata-kata seperti "bisa", "gagal", "tidak", "susah", dan "mantap" adalah fitur paling berpengaruh dalam membedakan sentimen, di mana skor Chi-Square yang tinggi menandakan hubungan kuat antara kata tersebut dan label sentimen tertentu.

6. *Data Mining*

Dalam klasifikasi Naive Bayes, probabilitas rata-rata setiap variabel pada data *training* untuk setiap kelas dihitung dan digunakan untuk mengklasifikasikan data *testing* (Arifat et al., 2023). Pada tahap ini, data dibagi menjadi dua bagian, yaitu 80% dialokasikan sebagai data pelatihan untuk membangun model dan 20% sisanya digunakan sebagai data pengujian untuk mengukur kinerja model. Setelah proses pembagian selesai, evaluasi dilakukan pada tahap *pattern evaluation* dengan memanfaatkan *confusion matrix* guna menilai tingkat akurasi dan performa model yang telah dibentuk.

7. *Pattern Evaluation*

Tahap ini menghasilkan evaluasi kinerja model yang dianalisis menggunakan metode *confusion matrix*, dengan hasil yang disajikan dalam bentuk tabel berisi metrik evaluasi seperti akurasi, presisi, *recall*, dan *F1-score* sebagai berikut.

Tabel 11. Perbandingan Hasil Evaluasi

Evaluasi	Naive Bayes	Naive Bayes Chi-Square
Akurasi	0.94	0.95
Presisi	0.98	0.98
Recall	0.94	0.96
F1-Score	0.96	0.97

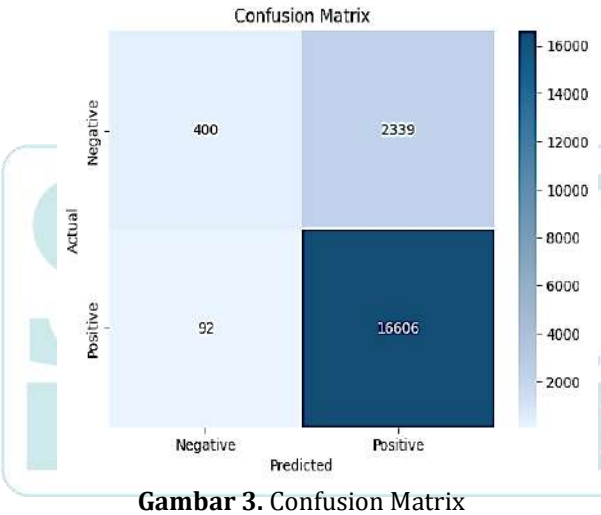
Berdasarkan Tabel 11, penambahan seleksi fitur Chi-Square memberikan peningkatan kinerja model. Secara keseluruhan, seleksi fitur Chi-Square

terbukti meningkatkan performa Naive Bayes meskipun perbedaannya relatif kecil, tetapi tetap menunjukkan model yang lebih efisien dan akurat.

Tabel 12. Hasil Evaluasi Perkelas

Evaluasi	Kelas Positif	Kelas Negatif
Presisi	0.88	0.81
Recall	0.99	0.15

Berdasarkan Tabel 12, diperoleh performa yang sangat baik pada kelas positif dengan presisi 0,88 dan *recall* 0,99. Namun, performa pada kelas negatif rendah, terutama recall yang hanya 0,15, menunjukkan bahwa sebagian besar ulasan negatif gagal dikenali dan banyak diklasifikasikan sebagai positif.



Gambar 3. Confusion Matrix

Tabel 13. *Confusion Matrix*

		Kelas Sebenarnya	
		Positif	Negatif
Kelas Prediksi	Positif	16606	2339
	Negatif	92	400

Berdasarkan Gambar 3. dan Tabel 13. *Confusion matrix* menunjukkan bahwa model sangat baik dalam mengklasifikasikan ulasan positif (TP = 16.606) dengan sedikit kesalahan prediksi (FN = 92), sehingga *recall* menjadi sangat tinggi. Namun, banyaknya ulasan negatif yang salah diklasifikasikan sebagai positif (FP = 2.339) menunjukkan bahwa model masih lemah dalam mengenali kelas negatif. Kondisi ini sejalan dengan hasil evaluasi sebelumnya dan mengindikasikan bias akibat ketidakseimbangan data, di mana dominasi ulasan positif membuat performa terhadap kelas negatif menurun. Ketidakseimbangan data yang ekstrem ini membuat performa model terhadap kelas minoritas (negatif) menjadi menurun.

8. *Knowledge Presentation*

Tahap ini menghasilkan visualisasi dari data penelitian guna memberikan gambaran umum atau informasi terkait ulasan pengguna aplikasi BRImo. Berikut merupakan gambar *wordcloud*.

Gambar 4. *Wordcloud* Ulasan Positif Aplikasi BRImo

Pada Gambar 4. Dominasi kata-kata seperti 'sangat', 'membantu', 'mudah', dan 'bagus' dalam visualisasi *wordcloud* menunjukkan bahwa kemudahan penggunaan dan manfaat aplikasi menjadi aspek utama yang diapresiasi oleh pelanggan. Ukuran kata yang lebih besar merepresentasikan frekuensi kemunculan yang tinggi, yang mencerminkan konsistensi dalam kepuasan pengguna terhadap pengalaman positif mereka dengan BRImo.

Gambar 5. *Wordcloud* Ulasan Negatif Aplikasi BRImo

Pada Gambar 5. Terlihat bahwa *wordcloud* ulasan negatif menampilkan kata

dominan seperti "tidak bisa" mengindikasikan keluhan pengguna yang sering mengalami kendala dalam menggunakan BRImo.

PENUTUP

Kesimpulan

Penelitian ini membuktikan bahwa kombinasi Naive Bayes dan Chi-Square efektif menganalisis sentimen pengguna BRImo. Kata kunci seperti "gagal" dan "tidak bisa" yang mengungkap fokus keluhan pada masalah teknis. Namun, model memiliki keterbatasan dalam mendeteksi sentimen negatif secara akurat akibat ketidakseimbangan data yang didominasi ulasan positif serta potensi ketidaksesuaian antara rating dan isi ulasan. Temuan ini menegaskan perlunya fokus perbaikan pada isu fungsional seperti akses dan kegagalan sistem, serta penggunaan teknik penanganan data tidak seimbang untuk meningkatkan akurasi.

Saran

Penerapan *cross- validation* direkomendasikan untuk meningkatkan ketepatan dalam mengevaluasi kinerja model. Analisis juga dapat diperluas dengan membandingkan efektivitas berbagai algoritma klasifikasi seperti SVM dan *Random Forest*, guna menentukan metode yang paling unggul. Selain itu, metode dengan mempertimbangkan imbalance data bisa memberikan pemahaman yang lebih komprehensif terhadap pola sentimen pengguna.

DAFTAR PUSTAKA

Andriani, N., & Wibowo, A. (2021). Implementasi Text Mining Klasifikasi Topik Tugas Akhir Mahasiswa Teknik Informatika Menggunakan Pembobotan TF-IDF dan Metode Cosine Similarity Berbasis Web. *Senamika*, *September*, 130–137. https://conference.upnvj.ac.id/index.php/senamika/article/view/1807%0Ahttp

Anisah, S., Honggowibowo, A. S., & Pujiastuti, A. (2016). Klasifikasi Teks Menggunakan Chi Square Feature Selection Untuk Menentukan Komik Berdasarkan Periode, Materi Dan Fisikdengan Algoritma Naïve Bayes. *Compiler*, *5*(2). https://doi.org/10.28989/compiler.v5i2.171

Arifat, M., Putri, W. A., & Mufida, A. S. (2023). Penerapan Metode Naive Bayes Classifier Untuk Klasifikasi Indeks Pembangunan Manusia Di Provinsi Jawa Timur. *Jurnal Statistika Dan Komputasi*, *2*(1), 31–43. https://doi.org/10.32665/statkom.v2i1.1661

Arminda, N. F., Sulistiyowati, N., & Padilah, T.N. (2023). Implementasi Algoritma Multinomial Naive Bayes Pada Analisis Sentimen Terhadap Ulasan Pengguna Aplikasi Brimo. *JATI*, *7*(3), 1817–1822. https://doi.org/10.36040/jati.v7i3.7012

Chen, H., Hu, S., Hua, R., & Zhao, X. (2021). Improved naive Bayes classification algorithm for traffic risk management. *EURASIP Journal on Advances in Signal Processing*, 2021(1). https://doi.org/10.1186/s13634-021-00742-6

Elysa, N. S., Arini, L., Murad, D. F., & Leandros, R. (2023). User Experience Satisfaction Analysis of Customers on the BRI Mobile Application (BRImo). *Procedia Computer Science*, *227*, 680–689. https://doi.org/10.1016/j.procs.2023.10.572

Ernayanti, T., Mustafid, M., Rusgiyono, A., & Hakim, A. R. (2022). Penggunaan Seleksi Fitur Chi-Square Dan Algoritma Multinomial Naïve Bayes Untuk Analisis Sentimen Pelangggan Tokopedia. *Jurnal*

Gaussian, *11*(4), 562–571. https://doi.org/10.14710/j.gauss.11.4.562-571

Farissa, R. A., Mayasari, R., & Umaidah, Y. (2021). Perbandingan Algoritma K-Means dan K-Medoids Untuk Pengelompokkan Data Obat dengan Silhouette Coefficient di Puskesmas Karangsambung. *Journal of Applied Informatics and Computing*, *5*(2), 109–116. https://doi.org/10.30871/jaic.v5i1.3237

Gautam, P. K., & Waoo, A. A. (2024). Improved Hybrid Feature Selection Approach for Sentiment Classification : Integrating Chi-Square and Recursive Feature Elimination. *Frontiers in Health Informatics*. *13*(3), 10570–10582. https://healthinformaticsjournal.com/index.php/IJMI/article/view/1008

Habiba, A., Isnanto, R. R., & Suseno, J. E. (2023). The Effect of Chi Square Feature Selection on the Naïve Bayes Algorithm on the Analysis of Indonesian Society's Sentiment About Face-to-Face Learning During the Covid-19 Pandemic. *JST (Jurnal Sains Dan Teknologi)*, *12*(1), 190-199. https://doi.org/10.23887/jstundiksha.v12i1.51899

Insan, M. K. K., Hayati, U., & Nurdiawan, O. (2023). Analisis Sentimen Aplikasi Brimo Pada Ulasan Pengguna Di Google Play Menggunakan Algoritma Naive Bayes. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *7*(1), 478–483. https://doi.org/10.36040/jati.v7i1.6373

Paembonan, S., & Abduh, H. (2021). Penerapan Metode Silhouette Coefficient untuk Evaluasi Clustering Obat. *PENA TEKNIK: Jurnal Ilmiah Ilmu-Ilmu Teknik*, *6*(2), 48. https://doi.org/10.51557/pt_jiit.v6i2.659

Permana, I. S., Halim, C., Nenti, N.S., & Riza, N. (2022). Analisis Kinerja Keuangan Dengan Menggunakan Rasio Likuiditas, Solvabilitas Dan Profitabilitas Pada PT. Bank BNI (Persero), TBK. *Jurnal Aktiva Riset Akuntansi Dan Keuangan*, *4*(1), 32–43. https://doi.org/10.52005/aktiva.v4i1.150

Prastyo, P. H., Ardiyanto, I., & Hidayat, R. (2021). A Combination of Query Expansion Ranking and GA-SVM for Improving Indonesian Sentiment Classification Performance. *Procedia Computer Science*, *189*, 108–115. https://doi.org/10.1016/j.procs.2021.05.074

Ratiasasadara, P. W., Sudarno, S., & Tarno, T. (2022). Analisis Sentimen Penerapan Ppkm Pada Twitter Menggunakan Naive Bayes Classifier Dengan Seleksi Fitur Chi-Square. *Jurnal Gaussian*, *11*(4), 580–590. https://doi.org/10.14710/j.gauss.11.4.580-590

Rosyadi, A. C. N., Athoillah, M., & Fenny Fitriani. (2025). Analisis Sentimen Pengguna Twitter Mengenai Kotak Kosong Di Pilkada Indonesia Tahun 2024 Menggunakan Algoritma LSTM. *Komputika : Jurnal Sistem Komputer*, *14*(2), 193–202. https://doi.org/10.34010/komputika.v14i2.16976

Saepudin, S., Widiastuti, S., & Irawan, C. (2023). Sentiment Analysis of Social Media Platform Reviews Using the Naïve Bayes Classifier Algorithm. *Jurnal Sistem Informasi Dan Komputer*, *12*(2), 236–243. https://doi.org/10.32736/sisfokom.v12i2.1650

Sidabutar, G. V. (2023). TA : Analisis

Sentimen Opini Publik Terhadap Pelayanan BPJS Kesehatan Menggunakan Metode Improved K-Nearest Neighbor - Repositori Universitas Dinamika. *Dinamika.ac.id*. https://repository.dinamika.ac.id/id/eprint/7362/1/18410100067-2023-UNIVERSITASDINAMIKA.pdf

Wijaya, R. H., Marthasari, G. I., & Sri, C. (2021). Perbandingan Feature Selection Chi-Square Dan Query Expansion Ranking (QER) Pada Analisis Sentimen Terkait Revitalisasi Monas Menggunakan Metode Naïve Bayes Classifier - UMM Institutional Repository. *Umm.ac.id*. https://eprints.umm.ac.id/id/eprint/5943/13/Wijaya%20Marthasari%20Aditya%20-%20Analisis%20Sentimen%20Na%C3%AFve%20Bayes%20Seleksi%20Fitur%20Python.pdf