

ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI KROM BANK DIGITAL MENGGUNAKAN NAÏVE BAYES, LOGISTIC REGRESSION, DAN SUPPORT VECTOR MACHINE

Aldi Wahdana^{1*}, Nurdin², Muhammad Sayuti³

^{1,2,3}Universitas Malikussaleh; Kampus Utama Bukit Indah, Jl. Cot Teungku Nie, Reuleut, Muara Satu, Kota Lhokseumawe, Aceh 24355; (+62) 645 41373

Keywords:

Analisis Sentimen;
Krom Bank Digital;
Naive Bayes;
Logistic Regression, Support
Vector Machine.

Correspondent Email:

wahdanaaldi60@gmail.com

Abstrak. Perkembangan teknologi informasi mendorong meningkatnya penggunaan aplikasi perbankan digital yang menghasilkan banyak ulasan pengguna pada Google Play Store. Ulasan tersebut penting untuk evaluasi layanan, namun jumlahnya yang besar menyulitkan analisis manual sehingga diperlukan pendekatan text mining otomatis. Penelitian ini bertujuan menganalisis sentimen ulasan pengguna aplikasi Krom Bank Digital serta membandingkan kinerja algoritma Naive Bayes, Logistic Regression, dan Support Vector Machine (SVM). Dataset yang digunakan berjumlah 11.718 ulasan, dan setelah proses pelabelan serta penghapusan data netral diperoleh 11.463 data. Tahapan penelitian meliputi preprocessing teks, ekstraksi fitur menggunakan TF-IDF, dan proses klasifikasi. Hasil penelitian menunjukkan bahwa SVM memberikan performa terbaik dengan akurasi 95,60% dan cross validation 95,21%, dibandingkan dua algoritma lainnya. Dengan demikian, SVM menjadi model paling optimal dalam klasifikasi sentimen. Penelitian ini diharapkan dapat membantu pengembang aplikasi dalam meningkatkan kualitas layanan serta memberikan gambaran mengenai opini pengguna melalui analisis data ulasan berbasis machine learning.



Copyright © 2026 The Author(s),
Published by Jurnal Informatika dan
Teknik Elektro Terapan. This is an
open-access article distributed under
the terms of the Creative Commons
Attribution-NonCommercial 4.0
International License (CC BY-NC
4.0).

Abstract. The development of information technology has driven the increasing use of digital banking applications, which generate numerous user reviews on the Google Play Store. These reviews are crucial for service evaluation, but their large number complicates manual analysis, necessitating an automated text mining approach. This study aims to analyze the sentiment of Krom Bank Digital application user reviews and compare the performance of Naive Bayes, Logistic Regression, and Support Vector Machine (SVM) algorithms. The dataset used consisted of 11,718 reviews, and after labeling and removing neutral data, 11,463 were obtained. The research stages included text preprocessing, feature extraction using TF-IDF, and classification. The results showed that SVM provided the best performance with 95.60% accuracy and 95.21% cross-validation, compared to the other two algorithms. Thus, SVM is the most optimal model for sentiment classification. This research is expected to assist application developers in improving service quality and provide an overview of user opinions through machine learning-based review data analysis.

1. PENDAHULUAN

Perkembangan teknologi informasi telah mendorong peningkatan penggunaan aplikasi mobile dalam berbagai bidang, termasuk sektor perbankan digital [1]. Google Play Store menyediakan fitur penilaian dan ulasan yang memungkinkan pengguna memberikan opini terhadap aplikasi yang digunakan. Ulasan pengguna merupakan sumber informasi penting yang dapat digunakan untuk mengetahui tingkat kepuasan maupun ketidakpuasan pengguna terhadap suatu aplikasi [2]. Namun, jumlah ulasan yang sangat banyak menyebabkan proses analisis secara manual menjadi sulit dilakukan sehingga diperlukan metode otomatis untuk mengolah data teks tersebut [3].

Analisis sentimen merupakan salah satu cabang text mining yang digunakan untuk mengidentifikasi opini pengguna berdasarkan data teks [4]. Text mining merupakan bagian dari data mining yang digunakan untuk menemukan pola dan informasi penting dari dokumen teks secara otomatis [5]. Dalam implementasinya, analisis sentimen banyak digunakan untuk mengelompokkan opini pengguna ke dalam kategori sentimen positif dan negatif sehingga dapat membantu perusahaan dalam melakukan evaluasi terhadap layanan yang diberikan [6].

Krom Bank Digital merupakan salah satu aplikasi perbankan digital yang menyediakan berbagai layanan transaksi keuangan melalui smartphone. Seiring meningkatnya jumlah pengguna aplikasi Krom Bank Digital, jumlah ulasan pengguna pada Google Play Store juga mengalami peningkatan secara signifikan. Berdasarkan pengamatan terhadap ulasan pengguna, sebagian besar pengguna memberikan tanggapan positif terkait kemudahan penggunaan aplikasi dan kecepatan transaksi. Namun demikian, masih ditemukan berbagai keluhan seperti aplikasi error, gagal login, serta gangguan pada beberapa fitur layanan. Oleh karena itu, diperlukan analisis terhadap ulasan pengguna untuk mengetahui opini pengguna terhadap aplikasi Krom Bank Digital secara lebih terstruktur [2].

Dalam penelitian ini digunakan beberapa algoritma machine learning yaitu Naive Bayes, Logistic Regression, dan Support Vector Machine (SVM). Naive Bayes merupakan algoritma klasifikasi probabilistik yang sering digunakan pada klasifikasi teks karena

memiliki proses komputasi yang sederhana dan cepat [7][8]. Logistic Regression merupakan metode klasifikasi yang digunakan untuk menentukan probabilitas suatu kelas berdasarkan data input [9]. Sementara itu, Support Vector Machine dikenal memiliki performa yang baik dalam menangani data teks berdimensi tinggi sehingga banyak digunakan pada analisis sentimen [10] [11] [12].

Penelitian sebelumnya yang dilakukan oleh peneliti pada aplikasi Livin' by Mandiri menggunakan metode Support Vector Machine dan TF-IDF menunjukkan bahwa metode tersebut mampu menghasilkan performa klasifikasi sentimen yang baik pada data ulasan aplikasi mobile banking [1]. Penelitian lain mengenai analisis sentimen aplikasi bank digital menggunakan algoritma Naive Bayes menunjukkan bahwa metode tersebut mampu melakukan klasifikasi sentimen pengguna dengan tingkat akurasi yang cukup tinggi [3]. Selain itu, penelitian benchmarking beberapa algoritma machine learning menunjukkan bahwa Support Vector Machine memiliki performa lebih baik dibandingkan Logistic Regression dan Naive Bayes pada klasifikasi data teks berdimensi tinggi [10].

Namun demikian, penelitian yang secara khusus membandingkan algoritma Naive Bayes, Logistic Regression, dan Support Vector Machine pada ulasan pengguna aplikasi Krom Bank Digital masih terbatas. Oleh karena itu, penelitian ini dilakukan untuk membandingkan performa ketiga algoritma tersebut dalam klasifikasi sentimen ulasan pengguna Krom Bank Digital menggunakan metode TF-IDF.

Sebelum proses klasifikasi dilakukan, data teks terlebih dahulu melalui tahapan preprocessing seperti case folding, cleaning, tokenizing, stopword removal, dan stemming. Selanjutnya dilakukan ekstraksi fitur menggunakan metode TF-IDF untuk mengubah data teks menjadi representasi numerik yang dapat diproses oleh algoritma machine learning [13].

Tujuan penelitian ini adalah membandingkan performa algoritma Naive Bayes, Logistic Regression, dan Support Vector Machine dalam klasifikasi sentimen ulasan pengguna Krom Bank Digital. Selain itu, penelitian ini juga bertujuan untuk mengetahui algoritma terbaik dalam menganalisis opini pengguna berdasarkan data ulasan pada Google

Play Store sehingga dapat membantu pengembang aplikasi dalam meningkatkan kualitas layanan aplikasi Krom Bank Digital.

2. TINJAUAN PUSTAKA

2.1 Analisis Sentimen

Analisis sentimen merupakan teknik dalam Natural Language Processing (NLP) yang digunakan untuk mengidentifikasi dan mengklasifikasikan opini atau perasaan seseorang terhadap suatu topik ke dalam kategori positif, negatif, atau netral [14]. Analisis sentimen telah banyak diterapkan pada berbagai bidang, seperti analisis ulasan produk, evaluasi layanan, serta pengukuran opini masyarakat terhadap kebijakan atau isu tertentu melalui media digital [14]. Selain itu, analisis sentimen menjadi salah satu pendekatan yang efektif untuk memahami persepsi publik dengan memanfaatkan data teks yang diperoleh dari media sosial maupun platform digital lainnya [15].

2.2 Naïve Bayes

Naive Bayes merupakan algoritma klasifikasi berbasis probabilitas yang menggunakan Teorema Bayes untuk menentukan kelas suatu data [6]. Algoritma ini mengasumsikan bahwa setiap fitur bersifat independen sehingga proses klasifikasi menjadi lebih sederhana dan cepat. Persamaan Naive Bayes ditunjukkan sebagai berikut:

$$P(H | X) = \frac{P(X | H)P(H)}{P(X)}$$

Keterangan:

$P(H|X)$ = probabilitas kelas terhadap data,

$P(X|H)$ = probabilitas data terhadap kelas,

$P(H)$ = probabilitas awal kelas,

$P(X)$ = probabilitas data.

Contoh Perhitungan

Misalnya probabilitas sentimen positif sebesar 0,8 dan sentimen negatif sebesar 0,2.

$$0.8 > 0.2$$

Karena probabilitas positif lebih besar, maka ulasan diklasifikasikan sebagai sentimen positif.

2.3 Logistic regression

Logistic Regression merupakan algoritma klasifikasi yang digunakan untuk memprediksi probabilitas suatu data terhadap kelas tertentu [7]. Algoritma ini menggunakan fungsi sigmoid untuk menghasilkan nilai probabilitas antara 0 dan 1. Persamaan fungsi sigmoid ditunjukkan sebagai berikut:

$$f(x) = \frac{1}{1 + e^{-x}}$$

Keterangan:

$f(x)$ = probabilitas output,

e = bilangan eksponensial,

x = nilai input.

Contoh Perhitungan

Misalnya diketahui nilai input sebesar 2.

$$f(2) = \frac{1}{1 + e^{-2}} = 0.88$$

Karena nilai probabilitas lebih besar dari 0,5 maka data diklasifikasikan sebagai sentimen positif.

2.4 Support Vector Machine

Support Vector Machine (SVM) merupakan algoritma klasifikasi yang bekerja dengan mencari hyperplane terbaik untuk memisahkan data antar kelas [10]. SVM memiliki performa yang baik dalam klasifikasi data teks berdimensi tinggi. Persamaan hyperplane pada SVM ditunjukkan sebagai berikut:

$$w \cdot x + b = 0$$

Keterangan:

w = bobot hyperplane,

x = data input,

b = bias.

Contoh Perhitungan

Misalnya diketahui:

$$w = 2, x = 3, b = -4$$

Maka:

$$(2 \times 3) + (-4) = 2$$

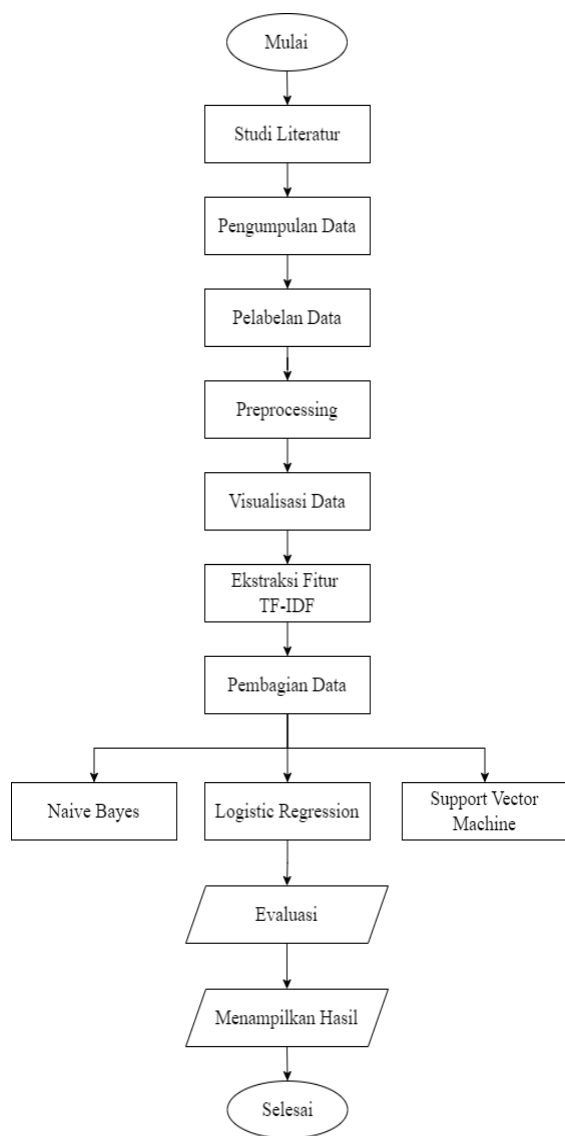
Karena hasil bernilai positif, maka data diklasifikasikan ke dalam sentimen positif.

3. METODE PENELITIAN

Metodologi penelitian merupakan tahapan yang dilakukan secara sistematis dalam proses analisis sentimen ulasan pengguna Krom Bank Digital menggunakan metode Naive Bayes, Logistic Regression, dan Support Vector Machine.

3.1 Tahapan Penelitian

Alur tahapan penelitian dilakukan secara sistematis mulai dari proses awal hingga tahap akhir penelitian. Diagram penelitian ini digunakan untuk memberikan gambaran mengenai urutan tahapan dan proses yang diterapkan selama pelaksanaan penelitian. Berikut ini alur diagram penelitian yang disajikan dalam bentuk flowchart.



Gambar 1. Flowchart Alur Penelitian

3.2. Studi Literatur

Tahap ini dilakukan dengan mempelajari jurnal, buku, dan penelitian terdahulu yang berkaitan dengan analisis sentimen, text mining, TF-IDF, serta algoritma Naive Bayes, Logistic Regression, dan Support Vector Machine (SVM) [1].

3.3. Pengumpulan Data

Data penelitian diperoleh dari ulasan pengguna aplikasi Krom Bank Digital pada Google Play Store dengan cara scraping data ulasan 3 tahun terakhir yaitu 2024-2026. Dataset yang digunakan berjumlah 11.718 data ulasan pengguna.

3.4. Pelabelan Data

Pelabelan data dilakukan berdasarkan rating yang diberikan oleh pengguna. Rating 4 dan 5 dikategorikan sebagai sentimen positif, sedangkan rating 1 dan 2 dikategorikan sebagai sentimen negatif. Rating 3 dikategorikan sebagai sentimen netral dan tidak digunakan dalam proses penelitian agar hasil klasifikasi lebih optimal [2].

3.5. Preprocessing

Tahap preprocessing dilakukan untuk membersihkan dan mempersiapkan data teks sebelum proses klasifikasi. Tahapan preprocessing meliputi:

1. Case Folding, yaitu mengubah seluruh huruf menjadi huruf kecil.
2. Cleaning, yaitu menghapus tanda baca, angka, URL, simbol, dan karakter yang tidak diperlukan.
3. Tokenizing, yaitu memecah kalimat menjadi kata-kata.
4. Stopword Removal, yaitu menghapus kata-kata umum yang tidak memiliki makna penting.
5. Stemming, yaitu mengubah kata menjadi bentuk dasar.

Tahap ini bertujuan agar data teks menjadi lebih terstruktur dan mudah diproses oleh algoritma machine learning.

3.6. Visualisasi Data

Visualisasi data dilakukan menggunakan WordCloud untuk menampilkan kata-kata yang paling sering muncul pada data ulasan pengguna. Visualisasi ini bertujuan untuk memberikan gambaran umum mengenai opini pengguna terhadap aplikasi Krom Bank Digital.

3.7. TF-IDF

Tahap TF-IDF (Term Frequency-Inverse Document Frequency) dilakukan untuk mengubah data teks menjadi representasi numerik berdasarkan bobot kemunculan kata dalam dokumen [2]. Hasil dari proses TF-IDF

digunakan sebagai fitur input pada proses klasifikasi.

3.8. Pembagian Data

Data yang telah diproses kemudian dibagi menjadi data training dan data testing menggunakan metode train-test split. Pada penelitian ini digunakan pembagian data sebesar 80% untuk training dan 20% untuk testing agar model dapat dilatih dan diuji secara optimal.

3.9. Klasifikasi

Tahap klasifikasi dilakukan menggunakan tiga algoritma machine learning yaitu Naive Bayes, Logistic Regression, dan Support Vector Machine (SVM). Ketiga algoritma tersebut digunakan untuk mengklasifikasikan data ulasan ke dalam kategori sentimen positif dan negatif [16].

3.10. Evaluasi

Tahap evaluasi dilakukan untuk mengetahui performa model klasifikasi. Evaluasi dilakukan menggunakan accuracy, precision, recall, f1-score, confusion matrix, dan cross validation [17]. Tahap ini bertujuan untuk mengetahui tingkat akurasi dan kestabilan model dalam melakukan klasifikasi sentimen.

3.11. Menampilkan Hasil

Tahap terakhir dilakukan dengan menampilkan hasil penelitian berupa nilai akurasi masing-masing algoritma, confusion matrix, classification report, serta grafik perbandingan performa model. Dari hasil tersebut kemudian ditentukan algoritma terbaik dalam klasifikasi sentimen ulasan pengguna Krom Bank Digital.

4. HASIL DAN PEMBAHASAN

Pada penelitian ini dilakukan analisis sentimen terhadap ulasan pengguna aplikasi Krom Bank Digital menggunakan algoritma Naive Bayes, Logistic Regression, dan Support Vector Machine (SVM). Proses penelitian meliputi pengumpulan data menggunakan teknik scraping, preprocessing, ekstraksi fitur TF-IDF, klasifikasi, dan evaluasi model untuk mengetahui algoritma dengan performa terbaik dalam klasifikasi sentimen pengguna.

4.1. Hasil Pengumpulan Data

Data penelitian diperoleh dari ulasan pengguna aplikasi Krom Bank Digital pada Google Play Store periode tahun 2024 hingga 2026. Proses pengumpulan data dilakukan menggunakan teknik scraping data ulasan sehingga diperoleh sebanyak 11.718 data ulasan pengguna. Contoh hasil scraping data ulasan dapat dilihat pada Tabel 1.

Tabel 1. Contoh Hasil Scrapping Data

Username	Rating	waktu	Ulasan
Pengguna 1	4	15/5/2026	Mantap
Pengguna 2	5	15/5/2026	bagussss sss
Pengguna 3	3	15/5/2026	bunga deposito nya bwsar
Pengguna 4	2	15/5/2026	No cs GK aktip2 parah...
Pengguna 5	4	15/5/2026	bgss dann terpercaya

Setelah proses pengumpulan data selesai, dilakukan pengecekan terhadap data kosong pada kolom ulasan. Data yang tidak memiliki isi ulasan dihapus agar tidak mempengaruhi proses klasifikasi. Selanjutnya dilakukan proses pelabelan sentimen berdasarkan rating pengguna. Rating 4 dan 5 dikategorikan sebagai sentimen positif, sedangkan rating 1 dan 2 dikategorikan sebagai sentimen negatif. Rating 3 dikategorikan sebagai sentimen netral dan tidak digunakan dalam penelitian.

Setelah dilakukan penghapusan data sentimen netral, jumlah dataset akhir yang digunakan sebanyak 11.463 data. Distribusi sentimen dataset dapat dilihat pada Tabel 2.

Tabel 2. Distribusi Sentimen Dataset

Sentimen	Jumlah Data
Positif	10.577
Negatif	886
Total	11463

Berdasarkan Tabel 2 dapat diketahui bahwa jumlah sentimen positif lebih dominan dibandingkan sentimen negatif. Hal ini menunjukkan bahwa sebagian besar pengguna memberikan tanggapan positif terhadap aplikasi Krom Bank Digital. Namun demikian, masih ditemukan ulasan negatif terkait error aplikasi, gagal login, dan gangguan transaksi.

4.2. Hasil Preprocessing Data

Tahap preprocessing dilakukan untuk membersihkan data teks sebelum proses klasifikasi dilakukan. Tahapan preprocessing yang digunakan meliputi case folding, cleaning, tokenizing, stopwords removal, dan stemming menggunakan Sastrawi. Tahap ini bertujuan agar data teks menjadi lebih bersih, terstruktur, dan mudah diproses oleh algoritma machine learning.

Tahap pertama yaitu case folding dilakukan dengan mengubah seluruh huruf menjadi huruf kecil (lowercase). Proses ini bertujuan untuk menyamakan bentuk penulisan kata sehingga tidak terjadi perbedaan akibat penggunaan huruf kapital. Hasil case folding dapat dilihat pada Tabel 3.

Tabel 3. Hasil Case Folding

Ulasan Asli	Hasil Case Folding
Aplikasi Sangat Bagus Dan Cepat	aplikasi sangat bagus dan cepat
Login Sering ERROR	login sering error
Transaksi Aman dan Mudah	transaksi aman dan mudah

Tahap berikutnya yaitu cleaning dilakukan untuk menghapus tanda baca, angka, URL, simbol, hashtag, mention, dan karakter lain yang tidak diperlukan. Hasil cleaning dapat dilihat pada Tabel 4.

Tabel 4. Hasil Cleaning

Ulasan	Hasil Cleaning
aplikasi bagus!!!	aplikasi bagus
login error!	login error
transaksi cepat 100%	transaksi cepat

Selanjutnya dilakukan proses tokenizing untuk memecah kalimat menjadi kumpulan kata

atau token. Hasil tokenizing dapat dilihat pada Tabel 5.

Tabel 5. Hasil Tokenizing

Ulasan	Hasil Tokenizing
aplikasi sangat bagus	[aplikasi, sangat, bagus]
login sering error	[login, sering, error]
transaksi cepat aman	[transaksi, cepat, aman]

Tahap berikutnya yaitu stopwords removal dilakukan untuk menghapus kata-kata umum yang tidak memiliki makna penting seperti “dan”, “yang”, “di”, dan “ke”. Hasil stopwords removal dapat dilihat pada Tabel 6.

Tabel 6. Hasil Stopword

Ulasan	Hasil Stopword Removal
aplikasi sangat bagus dan cepat	aplikasi bagus cepat
login yang sering error	login sering error
transaksi aman dan mudah	transaksi aman mudah

Tahap terakhir yaitu stemming dilakukan untuk mengubah kata menjadi bentuk dasar menggunakan library Sastrawi. Hasil dapat dilihat pada Tabel 7.

Tabel 7. Hasil Stemming

Ulasan	Hasil Stemming
aplikasi membantu pengguna	aplikasi bantu guna
login mengalami kesalahan	login alami salah
transaksi berjalan cepat	transaksi jalan cepat

Berdasarkan hasil preprocessing, data teks menjadi lebih bersih dan terstruktur sehingga dapat meningkatkan kualitas fitur pada proses klasifikasi sentimen.

4.3. Hasil Visualisasi WordCloud

Visualisasi WordCloud digunakan untuk mengetahui kata-kata yang paling sering muncul pada data ulasan pengguna.

Berdasarkan hasil visualisasi, kata-kata yang dominan muncul antara lain “aplikasi”, “bagus”, “mudah”, “mantap”, dan “krom”. Berikut ini hasil visualisasi kata dominan disajikan dalam bentuk gambar, dapat dilihat pada gambar 2.



Gambar 2. Hasil Visualisasi WordCloud

4.4 Hasil TF-IDF dan Pembagian Data

Tahap ekstraksi fitur dilakukan menggunakan metode TF-IDF dengan kombinasi unigram dan bigram serta maksimal 5000 fitur menggunakan library TF-IDF Vectorizer. Pada tahap ini, data hasil preprocessing dan stemming diubah menjadi representasi numerik berdasarkan bobot kata yang muncul pada setiap dokumen ulasan. Menggunakan Source code berikut ini.

```
tfidf = TfidfVectorizer(
    max_features=5000,
    ngram_range=(1,2)
)
```

Gambar 3. Source Code TF-IDF

data training dan data testing menggunakan rasio 80:20.

Hasil pembagian data dapat dilihat pada Tabel 8.

Tabel 8. Pembagian Data

Jenis Data	Jumlah
Data Training	9.170
Data Testing	2.293

Data training digunakan untuk melatih model klasifikasi, sedangkan data testing

digunakan untuk menguji performa model terhadap data baru.

4.5 Hasil Pengujian Naïve Bayes

Pengujian menggunakan algoritma Naive Bayes menghasilkan akurasi sebesar 94,46% dengan rata-rata cross validation sebesar 93,91%.

Tabel 9. Hasil Pengujian Naive Bayes

Accuracy	Precision Negatif	Recall Negatif	F1- Score negatif
94.46 %	0.84	0.35	0.49

Hasil pengujian menunjukkan bahwa Naive Bayes memiliki performa yang cukup baik dalam klasifikasi sentimen positif. Namun, model masih mengalami kesulitan dalam mengenali sentimen negatif karena nilai recall pada kelas negatif masih rendah.

4.6 Hasil Pengujian Logistic Regression

Pengujian menggunakan algoritma Logistic Regression menghasilkan akurasi sebesar 94,24% dengan rata-rata cross validation sebesar 94,25%.

Tabel 10. Hasil Pengujian Logistic Regression

Accuracy	Precision Negatif	Recall Negatif	F1- Score negatif
94.24%	0.88	0.29	0.44

Berdasarkan hasil pengujian, Logistic Regression memiliki performa yang cukup stabil, namun kemampuan model dalam mengenali sentimen negatif masih kurang optimal akibat distribusi data yang tidak seimbang.

4.7 Hasil Pengujian SVM

Pengujian menggunakan algoritma Support Vector Machine menghasilkan performa terbaik dengan akurasi sebesar 95,60% dan rata-rata cross validation sebesar 95.21%.

Tabel 11. Hasil Pengujian SVM

Accuracy	Precision Negatif	Recall Negatif	F1- Score negatif
95,60%	0,78	0,60	0,68

Hasil pengujian menunjukkan bahwa SVM lebih baik dalam mengenali sentimen negatif dibandingkan model lainnya. Selain memiliki akurasi tertinggi, nilai recall dan f1-score pada kelas negatif juga lebih baik dibandingkan Naive Bayes dan Logistic Regression.

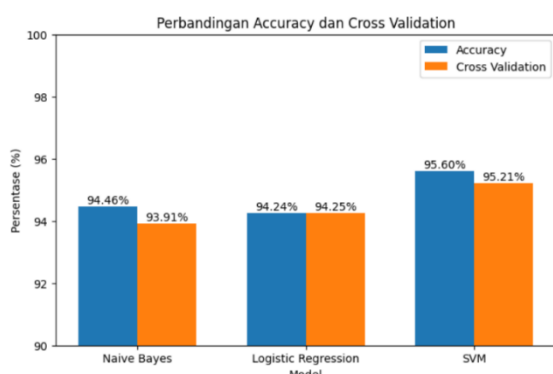
4.8 Perbandingan Performa Model

Perbandingan performa model dilakukan menggunakan metrik accuracy, precision, recall, f1-score, dan cross validation untuk mengetahui algoritma terbaik dalam klasifikasi sentimen.

Tabel 12. Accuracy dan Cross Validation

Model	Accuracy	Cross Validation
Naive Bayes	94.46%	93.91%
Logistic Regression	94.24%	94.25%
SVM	95.60%	95.21%

Berdasarkan Tabel 12, Support Vector Machine (SVM) memperoleh nilai accuracy dan cross validation tertinggi dibandingkan model lainnya. Berikut ini disajikan dalam bentuk barchart.

**Gambar 4.** Barchart Perbandingan Accuracy dan Cross Validation

Tabel berikutnya merupakan perbandingan Precision, Recall, dan F1-Score model.

Tabel 13. Precision dan Recall

Model	Precision	Recall
Naive Bayes	0.84	0.35
Logistic Regression	0.88	0.29
SVM	0.78	0.60

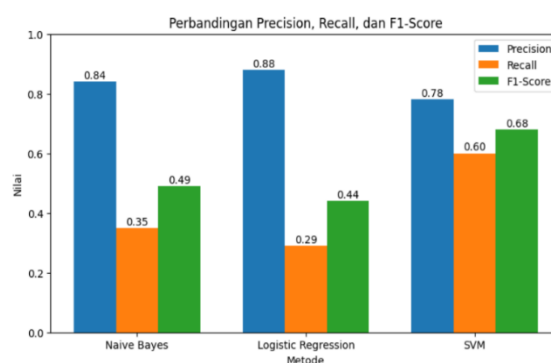
Berdasarkan Tabel 13, SVM memiliki nilai recall tertinggi sehingga lebih baik dalam mengenali data sentimen negatif dibandingkan model lainnya.

Tabel 14. F1-Score

Model	F1-Score
Naive Bayes	0.49
Logistic Regression	0.44
SVM	0.68

Berdasarkan Tabel 14, SVM menghasilkan F1-score tertinggi yang menunjukkan keseimbangan performa terbaik antara precision dan recall.

Berikut ini hasil perbandingan precision, recall, f1-score disajikan dalam bentuk barchart.

**Gambar 5.** Barchart Perbandingan Precision, Recall, F1-Score

5. KESIMPULAN

Berdasarkan hasil penelitian, analisis sentimen terhadap ulasan pengguna aplikasi Krom Bank Digital berhasil dilakukan menggunakan algoritma Naive Bayes, Logistic Regression, dan Support Vector Machine (SVM) melalui tahapan preprocessing serta ekstraksi fitur TF-IDF. Hasil pengujian

menunjukkan bahwa ketiga algoritma mampu memberikan performa klasifikasi yang baik dengan tingkat akurasi di atas 94%. Di antara ketiga metode tersebut, SVM memperoleh performa terbaik dengan nilai akurasi sebesar 95,60% dan cross validation sebesar 95,21%, sehingga dinilai paling optimal dalam mengklasifikasikan sentimen ulasan pengguna. Penelitian ini juga menunjukkan bahwa ketidakseimbangan distribusi data memengaruhi kemampuan model dalam mengenali sentimen negatif, khususnya pada metode Naive Bayes dan Logistic Regression.

Penelitian selanjutnya disarankan menerapkan teknik penyeimbangan data seperti SMOTE atau undersampling untuk meningkatkan kemampuan model dalam mengenali kelas minoritas. Selain itu, penggunaan metode ekstraksi fitur yang lebih kompleks seperti Word2Vec, FastText, atau BERT dapat dipertimbangkan guna menghasilkan representasi teks yang lebih baik. Pengembangan model berbasis deep learning, seperti LSTM maupun Transformer, juga dapat dilakukan untuk meningkatkan performa klasifikasi sentimen. Agar hasil penelitian lebih luas dan representatif, dataset sebaiknya diperoleh dari berbagai platform ulasan sehingga analisis tidak hanya terbatas pada satu sumber data.

DAFTAR PUSTAKA

- [1] A. Suharman and M. Kamayani Sulaeman, "Analisis Sentimen Pengguna Aplikasi Livin' by Mandiri Menggunakan Metode Support Vector Machine (SVM) dengan Ekstraksi Fitur TF-IDF dan Word2Vec," *Jurnal Pendidikan dan Teknologi Indonesia*, vol. 5, no. 8, pp. 2201–2212, Aug. 2025, doi: 10.52436/1.jpti.941.
- [2] M. Saputra and Sri Wahyuni, "ANALISIS SENTIMEN PENGGUNA PADA APLIKASI BANK DIGITAL KROM DENGAN ALGORITMA SUPPORT VECTOR MACHINE," *INFOTECH journal*, vol. 10, no. 2, pp. 327–332, Nov. 2024, doi: 10.31949/infotech.v10i2.11801.
- [3] A. Hilman Hakim *et al.*, "Analisis Sentimen Aplikasi Bank Digital Pada Google Play Store Menggunakan Algoritma Naive Bayes," *Journal of Informatics and Communications Technology*, vol. 7, no. 1, pp. 23–034, Jun. 2025, doi: 10.52661.
- [4] A. Setiawan,) Firman, N. Hasan, and R. Artikel, "ANALISIS SENTIMEN TANGGAPAN PENGGUNA APLIKASI BALE BY BTN MENGGUNAKAN METODE SUPPORT VECTOR MACHINE (SVM) 1) INFO ARTIKEL ABSTRAK," *Jurnal Ilmiah Teknik dan Ilmu Komputer*, vol. 4, no. 4, pp. 315–326, Nov. 2025, doi: 10.55123.
- [5] D. Mariyani, A. Irma Purnamasari, I. Ali, O. Nurdian, and R. Nurdian, "ANALISIS SENTIMEN ULASAN APLIKASI BANK JAGO MENGGUNAKAN SUPPORT VECTOR MACHINE DAN NEURAL NETWORK, SENTIMENT ANALYSIS OF BANK JAGO APPLICATION REVIEWS USING SUPPORT VECTOR MACHINE AND NEURAL NETWORK," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 14, no. 1, Jan. 2026, doi: 10.23960/jitet.v14i1.8775.
- [6] F. Adelia Irawan, A. Rialdy Atmadja, A. Wahana, T. Informatika, and U. Sunan Gunung Djati Bandung, "Analisis Sentimen Ulasan Aplikasi Bank Digital Menggunakan Algoritma Naive Bayes," 2024.
- [7] A. N. I. Pasha, E. F. Putri, L. Muthoharoh, A. Satria, and M. C. T. Manullang, "Hybrid TF--IDF Logistic Regression and MLP Neural Baseline for Indonesian Three-Class Sentiment Analysis on Social Media Text," *Cornell University*, vol. 6, no. 2, May 2026, doi: <https://doi.org/10.48550/arXiv.2605.07793>.
- [8] N. Nurdin, S. Abadi, and Y. Afrillia, "Classification System for Soil Types Suitable for Food Crops using Naïve Bayes Method," *Sistemasi: Jurnal Sistem Informasi*, vol. 13, no. 3, Apr. 2024, doi: DOI:<https://doi.org/10.32520/stmsi.v13i3.3956>.
- [9] T. Raisha Dwiva Putri, M. Rudi Sanjaya, M. Afriyan Firdaus, and D. Rosa indah, "Implementation of the K-Nearest Neighbors Algorithm for Sentiment Analysis on Aladin Bank Application Reviews," *Indonesian Journal of Machine Learning and Computer Science*, vol. 6, no. 2, pp. 608–618, Apr. 2026, doi: 10.57152/malcom.v6i2.2633.
- [10] N. Z. Zahra, S. Farhanatussaidah, N. N. Afifah, L. Muthoharoh, A. Satria, and M. C. T. Manullang, "Benchmarking Logistic Regression, SVM, Naive Bayes, and IndoBERT Fine-Tuning for Sentiment Analysis on Indonesian Product Reviews," *Cornell University*, vol. 4, no. 3, May 2026, doi: <https://doi.org/10.48550/arXiv.2605.03439>.
- [11] A. Alaiya, N. Nurdin, and C. Agusniar, "Sentiment Analysis of E-Commerce Product

- Reviews on Tokopedia Using Support Vector Machine,” *Journal of Applied Informatics and Computing (JAIC)*, vol. 9, no. 5, p. 2869, Oct. 2025, doi: <https://doi.org/10.30871/jaic.v9i5.10977>.
- [12] N. Nurdin, “Klasifikasi Penerima Bantuan dari Kepemilikan Kartu Pelaku Utama Sektor Kelautan dan Perikanan Dengan Metode Support Vector Machine,” *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 12, no. 3, Aug. 2024, doi: 10.23960/jitet.v12i3.4507.
- [13] S. F. Ngalo, I. R. H. T. Tangkawarow, and K. Santa, “Analisis Sentimen Terhadap Ulasan Aplikasi Bsgtouch Menggunakan Algoritma Support Vector Machine,” *Jurnal Minfo Polgan*, vol. 14, no. 2, pp. 2990–2998, Nov. 2025, doi: 10.33395/jmp.v14i2.15636.
- [14] G. hirzi Farizi, A. Voutama, and A. D. Hakim, “ANALISIS SENTIMEN MAHASISWA TERHADAP KETERSEDIAAN FASILITAS DI FAKULTAS ILMU KOMPUTER UNIVERSITAS SINGAPERBANGSA KARAWANG,” *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 13, no. 3, Jul. 2025, doi: 10.23960/jitet.v13i3.6544.
- [15] R. A. Munir, “ANALISIS SENTIMEN CUITAN DI MEDIA SOSIAL X TENTANG PROGRAM MAKAN BERGIZI GRATIS DENGAN METODE NLP,” *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 13, no. 3, Jul. 2025, doi: 10.23960/jitet.v13i3.6912.
- [16] D. Pebrianti, H. Kurniawan, L. Bayuaji, and R. Rusdah, “XgBoost Hyper-Parameter Tuning Using Particle Swarm Optimization for Stock Price Forecasting,” *Jurnal Ilmiah Teknik Elektro Komputer dan Informatika*, vol. 9, no. 4, pp. 1179–1195, Jan. 2024, doi: 10.26555/jiteki.v9i4.27712.
- [17] A. Dwi Syafutra, P. Korespondensi, and R. Artikel, “Analisis sentimen omnibus law di Twitter dengan machine learning dan teknik resampling,” *Jambura Journal of Informatics*, vol. 7, no. 1, pp. 22–35, 2025, doi: 10.37905/jji.v1i1.30395.