

PERBANDINGAN ALGORITMA MACHINE LEARNING UNTUK SENTIMEN ULASAN APLIKASI JOBSTREET

Fadhil Marzuqi¹, Yericco Purba², Thalut Syaputra³, Waeisul Bismi⁴, Ika Kurniawati⁵, Rizal Fahlap⁶

^{1,2,3,4,5,6}Universitas Bina Sarana Informatika, Jakarta, Indonesia

Responden: Waeisul.wbn@bsi.ac.id

ABSTRAK

Analisis sentiment digunakan untuk mengidentifikasi opini pengguna terhadap aplikasi digital berdasarkan teks ulasan. JobStreet sebagai platform rekrutmen yang populer memiliki ribuan ulasan pada *Google Play Store*, sehingga sesuai untuk dianalisis menggunakan pendekatan *machine learning*. Penelitian ini membandingkan empat algoritma—*Naïve Bayes*, *Support Vector Machine (SVM)*, *Logistic Regression*, dan *Random Forest*—untuk menentukan model terbaik dalam klasifikasi sentimen ulasan JobStreet. Sebanyak 5.000 ulasan dikumpulkan menggunakan *google-play-scrapers*, kemudian diproses melalui *preprocessing* yang meliputi *casefolding*, *tokenizing*, *stopword removal*, normalisasi, dan *stemming*. Setelah pembersihan, 3.846 ulasan valid direpresentasikan menggunakan *TF-IDF*. Evaluasi dilakukan dengan skema *train-test split* 80:20 dan *5-Fold Cross Validation*. Hasil menunjukkan bahwa *Naïve Bayes* memperoleh akurasi tertinggi 85.19%, diikuti *Logistic Regression* 84.42%, *Random Forest* 82.60%, dan *SVM* 81.69%. Hasil dalam memproses teks pendek. Penelitian ini memberikan gambaran komprehensif mengenai model terbaik untuk menganalisis sentimen aplikasi rekrutmen digital berbasis ulasan pengguna.

Kata Kunci: analisis sentimen, machine learning, JobStreet, TF-IDF, klasifikasi..

Riwayat Artikel :

Tanggal diterima : 09-10-2025

Tanggal revisi : 07-11-2025

Tanggal terbit : 01-12-2025

DOI :

<https://doi.org/10.31949/infotech.v11i2.16609>

INFOTECH journal by Informatika UNMA is licensed under CC BY-SA 4.0

Copyright © 2025 By Author



1. PENDAHULUAN

Aplikasi digital menyediakan ruang bagi pengguna untuk menyampaikan pengalaman dan persepsinya melalui ulasan platform seperti *Google Play Store*. Analisis sentiment telah digunakan dalam berbagai penelitian untuk menilai kualitas layanan aplikasi, seperti yang dilakukan pada aplikasi Shopee (Putri Anggraini & Winarsih, 2025), Pluang Bagas Akbar Maulana dkk. (2024), TikTok Sukirman dkk. (2024), aplikasi Korlantas (Nanda Ressaq Setiawan & Emil R. Kaburuan, 2023), serta aplikasi MyPertamina Gilbert Darmawan dkk. (2023). Selain itu, penelitian pada aplikasi JKN menunjukkan bahwa ulasan pengguna sering menyoroti kualitas layanan dan kendala teknis aplikasi, sehingga analisis sentimen menjadi penting untuk memahami persepsi pengguna terhadap aplikasi pelayanan publik (Tarwoto dkk., 2025) Pendekatan ini memungkinkan peneliti memahami kecenderungan opini baik positif maupun negatif.

Penelitian lain juga menunjukkan bahwa efektivitas model *machine learning* sangat dipengaruhi oleh karakteristik dataset. Implementasi algoritma *Naive Bayes* dan SVM pada aplikasi Canva oleh (Hariyadi dkk., 2024) membuktikan bahwa kualitas *preprocessing* dan representasi fitur memiliki dampak signifikan terhadap performa model. Selain itu studi (Alyaa Nadira dkk., 2023) mengungkapkan bahwa ulasan pengguna aplikasi keuangan umumnya mencerminkan kepuasan atau keluhan terhadap stabilitas sistem dan kemudahan penggunaan, sehingga analisis sentimen penting untuk mengevaluasi layanan berbasis aplikasi digital. Temuan serupa juga diperoleh pada penelitian aplikasi Ajaib (Nanda Dwi Kurniawan dkk., 2025), dimana variasi struktur teks pengguna berpengaruh terhadap performa model.

Sejumlah penelitian juga menunjukkan bahwa pemilihan algoritma *machine learning* sangat mempengaruhi kualitas klasifikasi sentimen. SVM memberikan kinerja kuat pada ulasan LinkedIn Gishella Septania Al-Husna dkk. (2024), sedangkan *Naive Bayes* terbukti unggul pada ulasan Shopee Riki Roska Rismansyah dkk. (2025) dan TikTokShop Octavia Salwa Dzaky Fadhillah dkk. (2025). *Logistic Regression* bekerja stabil pada ulasan aplikasi Netflix Nevita Cahaya Ramadani dkk. (2024), sedangkan *Random Forest* menunjukkan performansi konsisten pada analisis *Play Store* Muhammad Faiq Rahmatullah dkk. (2025) dan Apple Vision Pro (Rangga Rizky Pratama & Ryan Randy Suryono, 2025).

Di sisi lain, penelitian terbaru pada aplikasi Glints menunjukkan variasi opini yang berkaitan dengan masalah teknis pengguna Fanny Rahmasari dkk. (2025), dan studi lain membuktikan bahwa struktur teks pendek sangat mempengaruhi performa model Nanda Dwi Kurniawan dkk. (2025). Penelitian analisis sentimen juga banyak digunakan pada domain aplikasi karier seperti BYOND BSI (Firdaus Naifah Firzatullah & Nuroji, 2025).

Namun sampai saat ini, belum ada penelitian yang membandingkan secara langsung empat algoritma utama (*Naive Bayes*, SVM, *Logistic Regression*, dan *Random Forest*) secara bersamaan pada ulasan JobStreet di *Google Play Store*. Hal ini menjadi celah penelitian yang ingin diisi oleh studi ini.

Selama melakukan observasi awal terhadap ulasan pengguna JobStreet di *Google Play Store*, terlihat bahwa pola penulisan ulasan pada aplikasi karier memiliki karakteristik yang berbeda dibanding aplikasi layanan umum atau platform hiburan. Ulasan pengguna cenderung lebih ringkas, langsung pada inti permasalahan, dan sering berfokus pada pengalaman penggunaan aplikasi secara teknis, seperti kemudahan akses, kualitas antarmuka, serta kestabilan sistem. Ulasan panjang relatif jarang ditemukan, sementara keluhan teknis justru menjadi tema dominan yang muncul secara berulang.

Selain itu, pengguna yang meninggalkan ulasan negatif umumnya menuliskan keluhan yang terkait dengan proses masuk aplikasi, kendala notifikasi, keterlambatan pembaruan lowongan, dan masalah verifikasi akun. Sebaliknya, ulasan positif sering muncul dari pengguna yang merasa terbantu dengan kemudahan pencarian pekerjaan atau fitur pencarian yang sesuai dengan kebutuhan mereka. Pola ini menunjukkan bahwa sentimen pengguna JobStreet tidak hanya dipengaruhi oleh persepsi terhadap layanan inti, tetapi juga sangat bergantung pada stabilitas teknis dan pengalaman pengguna sehari-hari.

2. METODE

2.1 Pengumpulan Data

Data dikumpulkan menggunakan pustaka *google-play-scraper*, sama seperti yang dilakukan dalam studi Threads (Muhammad Arya Java dkk., 2024), Sirekap (Dede Ardian Tarigan dkk., 2025), dan aplikasi Korlantas (Nanda Ressaq Setiawan & Emil R. Kaburuan, 2023).

Menggunakan *google-play-scraper* telah terbukti menjadi cara yang paling efisien dan efektif untuk menangkap ulasan langsung dari *Google Play Store* dalam format yang terstruktur. Dengan cara ini, para peneliti dapat memperoleh sejumlah besar ulasan dan kami percaya bahwa ulasan yang dikumpulkan adalah pendapat pengguna yang autentik dan sah.

Dalam penelitian ini, 5.000 ulasan dikumpulkan sehingga dataset dapat memiliki keragaman pendapat yang cukup. Setiap ulasan berisi teks, penilaian, dan tanggal unggahan, yang kemudian menjadi dasar untuk pembersihan data dan analisis sentimen.

2.2 Preprocessing

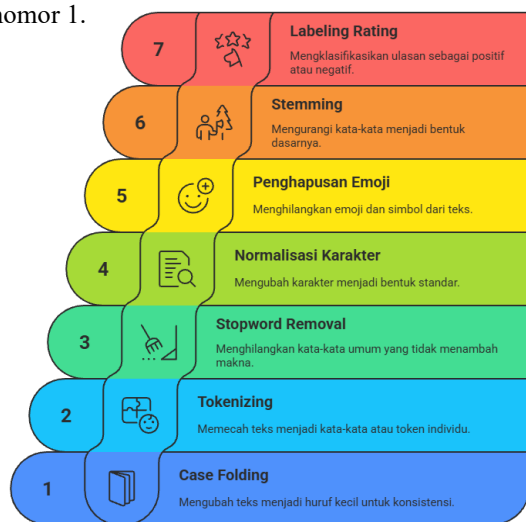
Tahapan pemrosesan mengadopsi metode penelitian sebelumnya pada Shopee (Putri Anggraini & Winarsih, 2025), TikTok (Sukirman dkk., 2024), dan Glints (Fanny Rahmasari dkk., 2025), Labeling sentimen dilakukan sesuai dengan pendekatan

(Satya Abdul Halim Bahtiar dkk., 2023) Rating 1-3 (negatif) dan 4-5 (positif).

Pemilihan skema pelabelan ini di dasarkan pada karakteristik ulasan aplikasi JobStreet yang cenderung bersifat ringkas dan *to the point*. Ulasan pada rating 3 pada umumnya tidak merepresentasikan sikap netral murni, melainkan lebih sering memuat keluhan ringan seperti keterlambatan notifikasi, proses login, atau ketidakpuasan terhadap fitur tertentu. Dengan kata lain, rating 3 menunjukkan ketidakpuasan tingkat rendah, bukan posisi tengah yang benar-benar seimbang. Apabila rating 3 diposisikan sebagai kategori netral, maka jumlah data netral akan menjadi sangat sedikit dan berpotensi menimbulkan ketidak seimbangan kelas yang dapat menurunkan performa model. Sedangkan jika rating 3 dihapus, maka hilang juga konteks opini pengguna yang sebenarnya bernilai informasi.

Oleh karena itu, rating 3 dimasukan ke dalam kelas negatif agar representasi sentimen lebih realistis dan distribusi data lebih stabil saat model dilatih.

Proses penelitian ini meliputi *casefolding*, *tokenizing*, *stopword removal*, normalisasi, *stemming*, dan *labeling rating*. Seperti gambar nomor 1.



Gambar 1. Proses Penelitian

2.3 Representasi Fitur

TF-IDF digunakan untuk konversi teks menjadi vektor numerik, sebagaimana diterapkan pada Canva (Hariyadi dkk., 2024), MyPertamina (Gilbert Darmawan dkk., 2023), dan BYOND BSI (Firdaus Naifah Firzatullah & Nuroji, 2025).

Dalam penelitian ini menggunakan *TF-IDF* dikarenakan dapat memberikan bobot pada kata-kata proporsional sesuai dengan frekuensi kemunculannya. Jika ada kata yang sering muncul pada satu ulasan, tetapi jarang muncul dalam ulasan lain, kata tersebut akan diberi bobot yang lebih tinggi. Selain itu, *TF-IDF* juga sangat sesuai untuk menganalisis ulasan karena umumnya bersifat pendek, langsung, dan mempunyai banyak variasi

kosakata. Algoritma ini akan mengubah teks yang abstrak menjadi sebuah vektor berdimensi tinggi yang dapat di proses oleh *machine learning*.

2.4 Algoritma Machine Learning

Penelitian ini menggunakan empat algoritma utama—*Naïve Bayes*, *Support Vector Machine (SVM)*, *Logistic Regression*, dan *Random Forest*—karena seluruhnya telah terbukti efektif dalam berbagai studi analisis sentimen berbasis ulasan aplikasi di *Google Play Store*.

Naïve Bayes dipilih karena performanya sangat baik pada teks pendek dan sederhana, sebagaimana ditunjukkan pada penelitian Shopee dan TikTokShop (Octavia Salwa Dzaky Fadhillah dkk., 2025; Riki Roska Rismansyah dkk., 2025), yang memiliki karakteristik serupa dengan ulasan JobStreet. Model probabilistik ini efisien dan mampu menangani variasi penulisan yang umum muncul pada ulasan pengguna.

SVM digunakan karena merupakan algoritma yang secara konsisten memberikan hasil unggul pada data berdimensi tinggi seperti *TF-IDF*. Penelitian (Gishella Septania Al-Husna dkk., 2024) dan (Nanda Ressq Setiawan & Emil R. Kaburuan, 2023) menunjukkan bahwa SVM mampu mengklasifikasikan ulasan aplikasi secara akurat, sehingga relevan sebagai perbandingan utama dalam penelitian ini.

Logistic Regression dipertimbangkan karena stabil dalam memproses data yang relatif seimbang dan sering memberikan performa kompetitif, sebagaimana ditunjukkan pada analisis sentimen aplikasi Netflix (Nevita Cahaya Ramadani dkk., 2024). Selain itu, model ini bekerja baik pada representasi fitur linier seperti *TF-IDF*.

Random Forest ditambahkan untuk melihat bagaimana pendekatan ensemble mengenai data yang memiliki banyak variasi kata dan *noise*. Penelitian (Muhammad Faiq Rahmatullah dkk., 2025) dan (Rangga Rizky Pratama & Ryan Randy Suryono, 2025) menunjukkan bahwa *Random Forest* cenderung memberikan performa yang stabil pada ulasan Play Store, sehingga relevan digunakan pada data JobStreet yang memiliki pola penulisan tidak seragam.

Keempat algoritma tersebut dipilih untuk memberikan gambaran komprehensif mengenai karakteristik performa model. *Naïve Bayes* populer digunakan untuk teks pendek dengan pola kata yang berulang. SVM dikenal kuat dalam memproses data berdimensi tinggi seperti representasi *TF-IDF*. *Logistic Regression* memberikan performa yang stabil serta interpretabilitas yang baik. Sementara itu, *Random Forest* mewakili pendekatan ensemble yang mampu menangani data dengan variasi kata yang luas.

Dengan demikian, keempat algoritma tersebut memberikan cakupan perbandingan yang memadai antara model probabilistik, linier, *margin-based*, dan

ensemble, sehingga memungkinkan penelitian menyeluruh terhadap efektivitas klasifikasi sentimen pada ulasan JobStreet.

2.5 Evaluasi

Evaluasi dilakukan menggunakan *accuracy*, *precision*, *recall*, dan *F1-score* sebagaimana pendekatan (Muhammad Faiq Rahmatullah dkk., 2025) dan (Dede Ardian Tarigan dkk., 2025), dengan *train-test split* 80:20 dan 5-Fold Cross Validation.

Kita bisa menggunakan *train-test split* sebagai cara untuk memeriksa kinerja model pada data-data yang belum dilatih model. Pada pembagian data di awal pelatihan model ini bisa di ambil untuk mengukur akurasi model. *5-Fold Cross Validation* digunakan untuk memverifikasi performa model yang konsisten terhadap beberapa partisi yang berbeda. Pendekatan ini penting untuk menghindari *overfit* dan untuk mendapatkan evaluasi yang lebih konsisten. Setiap metrik evaluasi bisa digunakan untuk memberikan evaluasi yang lebih komprehensif terlepas dari ketidakmerataan pembagian kelas.

3. PEMBAHASAN

bagian ini menyajikan hasil analisis sentimen ulasan JobStreet secara menyeluruh, meliputi distribusi data, pola kata, akurasi model, serta evaluasi Cross-Validation. Seluruh gambar digunakan sebagai bagian dari interpretasi hasil dan mendukung temuan penelitian.

3.1 Gambaran dan Distribusi Data

Penelitian ini menggunakan 5.000 ulasan mentah yang diperoleh melalui scraping Google Play Store. Proses pembersihan data dilakukan secara bertahap, dimulai dari pembuangan duplikat, data kosong, dan normalisasi teks. Hasilnya dapat dilihat pada gambar 2.

review	rating	date	app
0 mau verifikasi identitas rbel amat dan ternyata aplikasinya Indonesia Inggris, lebih ke aplikasi pencari kerja Inggris.	2	2025-11-13 15:41:33	JobStreet
1 apk terburuk yang pernah ada di muka bumi	1	2025-11-13 15:11:09	JobStreet
2 oke	5	2025-11-13 13:24:35	JobStreet
3 sangat membantu untuk mencari lowongan pekerjaan	5	2025-11-13 09:31:04	JobStreet
4 bagus	5	2025-11-13 07:19:50	JobStreet

Gambar 2. Contoh Data Mentah Sebelum

Pembersihan

Setelah pembersihan duplikat dan data kosong, diperoleh 3.887 ulasan sebagaimana ditunjukkan pada gambar 3.

review	rating	date	app	sentiment
0 mau verifikasi identitas rbel amat dan ternyata aplikasinya Indonesia Inggris, lebih ke aplikasi pencari kerja Inggris.	2	2025-11-13 15:41:33	JobStreet	negative
1 apk terburuk yang pernah ada di muka bumi	1	2025-11-13 15:11:09	JobStreet	negative
2 oke	5	2025-11-13 13:24:35	JobStreet	positive
3 sangat membantu untuk mencari lowongan pekerjaan	5	2025-11-13 09:31:04	JobStreet	positive
4 bagus	5	2025-11-13 07:19:50	JobStreet	positive

Gambar 3. Data Setelah Buang Duplikat & NaN

Distribusi sentimen sebelum *preprocessing* menunjukkan jumlah sentimen negatif sedikit lebih besar dibandingkan positif. Hal ini sejalan dengan temuan (Fanny Rahmasari dkk., 2025) mengenai dominannya keluhan teknis pada aplikasi lowongan kerja digital.

Setelah *preprocessing*, tersisa 3.846 data bersih sebagaimana terlihat pada gambar 4.

class	review	rating	sentiment
0	mau verifikasi identitas rbel amat dan ternyata aplikasinya Indonesia Inggris, lebih ke aplikasi pencari kerja Inggris.	2	negative
1	apk terburuk yang pernah ada di muka bumi	1	negative
2	oke	5	positive
3	sangat membantu untuk mencari lowongan pekerjaan	5	positive
4	bagus	5	positive

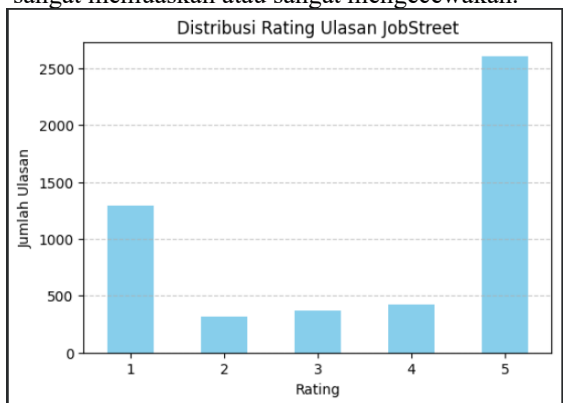
Gambar 4. Data Setelah *Preprocessing*

Perlu dicatat bahwa dalam penelitian ini, keseimbangan data tidak dilakukan melalui teknik *sampling* (*oversampling/undersampling*). Komposisi dataset terjadi secara alami, mengikuti proporsi asli yang muncul setelah proses pembersihan. Hal ini menyebabkan rating 1 dan 5 tetap mendominasi—sebuah pola yang umum pada aplikasi digital, dimana pengguna cenderung memberi ulasan saat sangat puas atau sangat kecewa. Dengan kata lain, meskipun terjadi pengurangan data pada tahap *cleaning*, perbedaan proporsi kelas tetap kuat dan secara signifikan berubah menjadi seimbang.

Keputusan untuk tidak melakukan teknik penyeimbangan seperti *SMOTE* atau *undersampling* dilakukan agar karakter asli opini pengguna tetap terjaga dan tidak menimbulkan distorsi pada perilaku ulasan yang sebenarnya. Dataset ini mencerminkan kondisi *real* pengguna JobStreet, sehingga hasil model dapat menggambarkan kecenderungan opini secara lebih natural dalam penggunaan nyata. Namun, struktur data ini juga menjadi salah satu faktor yang menyebabkan model seperti *SVM* dan *Random Forest* memiliki performa lebih rendah dibanding *Naïve Bayes* yang cenderung lebih adaptif terhadap pola kata yang repetitif dan tidak seimbang.

3.2 Distribusi Rating Ulasan

Gambar 5 menunjukkan bahwa rating 1 dan 5 mendominasi jumlah ulasan JobStreet. Pola distribusi ekstrim seperti ini umum terjadi pada aplikasi digital, dimana pengguna cenderung memberikan rating sangat tinggi atau sangat rendah. Fenomena serupa ditemukan pada penelitian (Fanny Rahmasari dkk., 2025) mengenai aplikasi Glints dan penelitian (Riki Roska Rismansyah dkk., 2025) pada aplikasi Shopee, dimana ulasan bintang 1 dan 5 mendominasi dibanding rating tengah. Hal ini menunjukkan bahwa pengguna lebih terdorong memberikan ulasan ketika memiliki pengalaman sangat memuaskan atau sangat mengecewakan.



Gambar 5. Distribusi Rating Ulasan JobStreet

Perubahan distribusi setelah pembersihan tetap mempertahankan karakter asli data, sehingga hasil klasifikasi mencerminkan kondisi *real* ulasan pengguna. Meskipun terdapat pergeseran proporsi akibat pembuangan duplikat dan data tidak valid, proses ini tidak ditunjukkan untuk menyeimbangkan kelas, melainkan menjaga kualitas data yang masuk ketahap pelatihan model. Dengan demikian, distribusi akhir dapat di anggap representatif dan tetap mencerminkan kecenderungan opini pengguna JobStreet secara natural.

3.3 Analisis WordCloud

Untuk memahami konteks kata domain dalam setiap sentimen, WordCloud digunakan sebagai visual eksploratif (Gambar 6).



Gambar 6. WordCloud Sentimen Positif dan Negatif

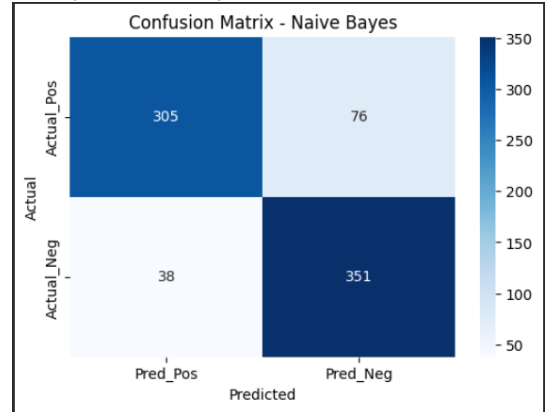
WordCloud positif menunjukan kata seperti *cari kerja, bagus, mudah, sangat bantu*, yang mengindikasikan kepuasan pengguna terhadap kemudahan mencari lowongan. Temuan ini sejalan dengan laporan (Bagas Akbar Maulana dkk., 2024) yang menunjukkan bahwa ulasan positif pada aplikasi karier domain berisi kemudahan navigasi dan rekomendasi lowongan.

WordCloud negatif didominasi kata seperti *masuk, lama, usahanya, update, login, error*. Ini mendukung argumen (Fanny Rahmasari dkk., 2025) bahwa aplikasi pekerjaan sering mendapatkan ulasan negatif terkait masalah teknis, bukan layanan inti.

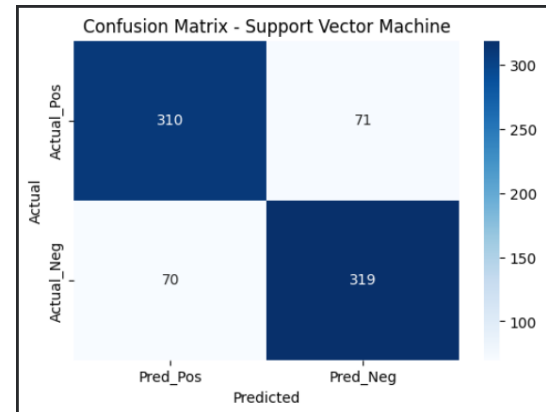
3.4 Evaluasi Model Menggunakan Confusion Matrix

Naïve Bayes memiliki nilai *accuracy* 85.19%, *precision* 85.53%, *recall* 85.19%, dan *F1-score* 85.15%. *Support Vector Machine* memiliki nilai *accuracy* 81.69%, *precision* 81.69%, *recall* 81.69%, dan *F1-score* 81.69%. *Logistic Regression* memiliki nilai *accuracy* 84.42%, *precision* 84.42%, *recall* 84.42%, dan *F1-score* 84.41%. *Random Forest* memiliki nilai *accuracy* 82.60%, *precision* 82.60%, *recall* 82.60%, dan *F1-score* 82.60%.

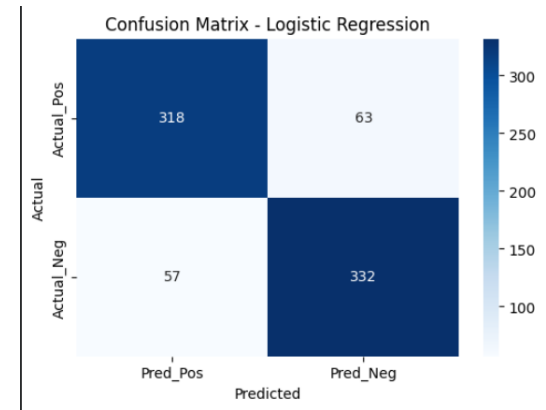
Evaluasi empat model dilakukan melalui confusion matrix (Gambar 7-10).



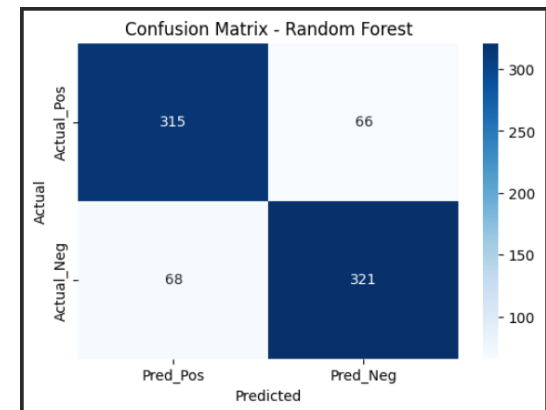
Gambar 7. Confusion Matrix-Naïve Bayes



Gambar 8. Confusion Matrix-Support Vector Machine



Gambar 9. Confusion Matrix-Logistic Regression



Gambar 10. Confusion Matrix-Random Forest

Dari keempat model, Naïve Bayes dan Logistic Regression terlihat memiliki prediksi benar yang paling tinggi, terutama pada kelas negatif. Hal ini menunjukkan bahwa model probabilistik dan linier bekerja baik pada pola kata yang repetitif (Nanda Ressaq Setiawan & Emil R. Kaburuan, 2023).

Sebaliknya, SVM menghasilkan misclassifications lebih besar. Berbeda dengan penelitian (Gishella Septania Al-Husna dkk., 2024), di mana SVM unggul pada ulasan LinkedIn, dataset JobStreet memiliki ulasan lebih pendek sehingga pemisahan margin optimal lebih sulit dicapai.

Sedangkan untuk Random Forest menunjukkan bahwa model ini mampu mengklasifikasikan kedua kelas dengan cukup baik, namun jumlah *false negative* yang dihasilkan lebih tinggi di banding Naïve Bayes dan Logistic Regression. Hal ini terjadi karena banyak fitur *TF-IDF* menyebabkan setiap pohon keputusan melihat informasi yang berbeda, sehingga konsistensi prediksi menurun (Rangga Rizky Pratama & Ryan Randy Suryono, 2025). Meskipun demikian, performanya tetap lebih stabil dibanding SVM dan masih Relevan sebagai model pembandingan dalam penelitian ini.

3.5 Train-Test Split

setelah dilakukan proses pembagian data menggunakan proporsi 80% train dan 20% test, diperoleh 3.076 data pada train set dan 770 data pada test set. Hasil di tunjukan pada gambar 11.

```
Train size: 3076
Test size : 770
Distribusi label train:
sentiment
negative 1556
positive 1520
Name: count, dtype: int64

Distribusi label test:
sentiment
negative 389
positive 381
Name: count, dtype: int64
```

Gambar 11. Train-Test Split

3.6 Evaluasi Cross-Validation

Untuk nilai kestabilan performa model, dilakukan 5-Fold Cross-Validation. Hasilnya ditampilkan pada Gambar 12.

Dari gambar dan tabel terlihat bahwa:

1. Logistic Regression (mean accuracy 0.8489) dan Naïve Bayes (0.8489) menjadi dua model paling stabil.
2. Naïve Bayes memiliki standar deviasi paling kecil (± 0.0076), menandakan konsistensi antarfold sangat tinggi.
3. SVM memiliki mean accuracy lebih rendah, sejalan dengan kinerjanya pada confusion matrix.
4. Random Forest memiliki fluktuasi terbesar, terkait tingginya dimensi fitur *TF-IDF* yang

memengaruhi struktur pohon (Rangga Rizky Pratama & Ryan Randy Suryono, 2025).

Model	CV_Mean_Accuracy	CV_STD
2 Logistic Regression	0.848935	0.014173
0 Naive Bayes	0.848934	0.007595
1 Support Vector Machine	0.833335	0.006463
3 Random Forest	0.830994	0.014773

Gambar 12. Hasil Cross-Validation empat model machine learning

Hasil ini menguatkan temuan sebelumnya bahwa model probabilistik dan linier paling cocok untuk dataset teks pendek dan padat (Putri Anggraini & Winarsih, 2025).

3.7 Komparasi Dengan Studi Terdahulu

hasil penelitian ini konsisten dengan pola yang ditemukan pada aplikasi Shopee (Riki Roska Rismansyah dkk., 2025) dan aplikasi Glints (Fanny Rahmasari dkk., 2025), di mana ulasan negatif didominasi masalah teknis dan model Naïve Bayes sering tampil sebagai model terbaik.

Namun hasil ini berbeda dengan penelitian (Gishella Septania Al-Husna dkk., 2024) dan (Nanda Ressaq Setiawan & Emil R. Kaburuan, 2023), di mana SVM mendominasi performa. Perbedaan ini menunjukkan bahwa efektivitas algoritma sangat dipengaruhi oleh panjang teks, repetisi kata, dan kompleksitas ulasan.

Temuan ini sejalan dengan penelitian (Riki Roska Rismansyah dkk., 2025) dan (Fanny Rahmasari dkk., 2025) yang menunjukkan bahwa efektivitas algoritma sangat bergantung pada struktur dan karakteristik dataset, sehingga tidak ada satu metode yang selalu unggul untuk semua jenis ulasan aplikasi.

4. KESIMPULAN

Penelitian ini menunjukkan bahwa Naïve Bayes merupakan algoritma terbaik untuk analisis sentimen ulasan JobStreet dengan akurasi 85.19%, diikuti Logistic Regression, Random Forest, dan SVM. Pola kata yang repetitif dan teks yang pendek membuat model probabilistik bekerja optimal. Proses *preprocessing* berkontribusi signifikan dalam meningkatkan kualitas data. WordCloud memperlihatkan bahwa keluhan teknis seperti login

dan error menjadi penyebab utama sentimen negatif. Hasil penelitian ini diharapkan menjadi masukan bagi pengembang JobStreet dan menjadi dasar penelitian lanjutan pada domain analisis sentimen aplikasi digital.

PUSTAKA

- Alyaa Nadira, Nanang Yudi Setiawan, & Welly Purnomo. (2023). *ANALISIS SENTIMEN PADA ULASAN APLIKASI MOBILE BANKING MENGGUNAKAN METODE NAÏVE BAYES DENGAN KAMUS INSET*.
- Bagas Akbar Maulana, Muhammad Jazilul Fahmi, Ari Muhamad Imran, & Nutriana Hidayati. (2024). Analisis Sentimen Terhadap Aplikasi Pluang Menggunakan Algoritma Naive Bayes dan Support Vector Machine (SVM). *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 4(2), 375–384. <https://doi.org/10.57152/malcom.v4i2.1206>
- Dede Ardian Tarigan, Zakarias Situmorang, & Rika Rosnelly. (2025). *ANALISIS SENTIMEN APLIKASI PLAYSTORE SIREKAP 2024 PASCA PILPRES DENGAN PERBANDINGAN METODE SUPPORT VECTOR MACHINE (SVM), NAÏVE BAYES CLASSIFIER DAN RANDOM FOREST*. 11(3), 661–670.
- Fanny Rahmasari, Nining Rahaningsih, Raditya Dinar Dana, & Cep Lukman Rohmat. (2025). OPTIMASI ANALISIS SENTIMEN APLIKASI GLINTS MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE (SVM). *Jurnal Informatika dan Teknik Elektro Terapan*, 13(1). <https://doi.org/10.23960/jitet.v13i1.5681>
- Firdaus Naifah Firzatullah, & Nuroji. (2025). *Analisis Sentimen Pengguna Aplikasi Beyond BSI Pada Google Play Store Menggunakan Algoritma SVM Dan Random Forest*. 9, 2025. <https://doi.org/10.47002/metik.v9i2.1089>
- Gilbert Darmawan, Syariful Alam, & M. Imam Sulisty. (2023). *ANALISIS SENTIMEN BERDASARKAN ULASAN PENGGUNA APLIKASI MYPERTAMINA PADA GOOGLE PLAYSTORE MENGGUNAKAN METODE NAÏVE BAYES INFO ARTIKEL ABSTRAK*. 2(3), 100–108. <https://doi.org/10.55123>
- Gishella Septania Al-Husna, Dian Asmarajati, Iman Ahmad Ihsanuddin, & Rina Mahmudati. (2024). *PERBANDINGAN METODE NAÏVE BAYES DAN SUPPORT VECTOR MACHINE UNTUK ANALISIS SENTIMEN PADA ULASAN PENGGUNA APLIKASI LINKEDIN 1) INFO ARTIKEL ABSTRAK*. 3(2), 139–144. <https://doi.org/10.55123>
- Hariyadi, Daud Firdo, & Muhammad Hadianur Al Rafi. (2024). Implementasi Algoritma Naïve Bayes dan Support Vector Machine pada Analisis Sentimen Ulasan Aplikasi Canva. *Jurnal Minfo Polgan*, 13(1), 261–269. <https://doi.org/10.33395/jmp.v13i1.13568>
- Muhammad Arya Java, Mohammad Syafrullah, Windarto, & Painem. (2024). Analisis Sentimen Ulasan Pengguna Aplikasi Threads pada Google Play Store Menggunakan Multinomial Naive Bayes dan Support Vector Machine. *Jurnal TICOM: Technology of Information and Communication*, 12(2). <https://github.com/nasalsabila/kamus-alay>
- Muhammad Faiq Rahmatullah, Poetri Lestari Lokapitasari Belluano, & Herdianti Darwis. (2025). Analisis Sentimen Review Aplikasi di Google Play Store Menggunakan Random Forest. *Literatur Informatika & Komputer*, 2(3), 380–389. <https://doi.org/10.33096/linier.v2i3.3149>
- Nanda Dwi Kurniawan, Praditya Rendi Ferdian, & Nurtriana Hidayati. (2025). Analisis Sentimen Algoritma Naïve Bayes, Support Vector Machine, dan Random Forest Pada Ulasan Aplikasi Ajaib. *Jurnal Nasional Teknologi dan Sistem Informasi*, 11(1), 87–97. <https://doi.org/10.25077/teknosi.v11i1.2025.87-97>
- Nanda Ressaq Setiawan, & Emil R. Kaburuan. (2023). Sentimen Analisis Review Aplikasi Digital Korlantas Pada Google Play Store Menggunakan Metode SVM. *Jurnal Sisfokom (Sistem Informasi dan Komputer)*, 12(1), 105–116. <https://doi.org/10.32736/sisfokom.v12i1.1614>
- Nevita Cahaya Ramadani, Imam Tahyudin, & Azhari Shouni Barkah. (2024). Perbandingan Algoritma Support Vector Machine, Decision Tree, dan Logistic Regresion Pada Analisis Sentimen Ulasan Aplikasi Netflix. *Jurnal Nasional Teknologi dan Sistem Informasi*, 10(2), 110–117. <https://doi.org/10.25077/teknosi.v10i2.2024.110-117>
- Octavia Salwa Dzaky Fadhillah, Jajam Haerul Jaman, & Carudin. (2025). PERBANDINGAN NAIVE BAYES, SUPPORT VECTOR MACHINE, LOGISTIC REGRESSION DAN RANDOM FOREST DALAM MENGANALISIS SENTIMEN MENGENAI TIKTOKSHOP. *Jurnal*

- Informatika dan Teknik Elektro Terapan*, 13(1).
<https://doi.org/10.23960/jitet.v13i1.5746>
- Putri Anggraini, & Winarsih. (2025). KOMPARASI NAÏVE BAYES, SUPPORT VECTOR MACHINE, DAN RANDOM FOREST DALAM ANALISIS SENTIMEN APLIKASI SHOPEE DI GOOGLE PLAY STORE. Dalam *Jurnal Mahasiswa Teknik Informatika* (Vol. 9, Nomor 3).
- Rangga Rizky Pratama, & Ryan Randy Suryono. (2025). PERFORMANCE COMPARISON OF NAIVE BAYES, SUPPORT VECTOR MACHINE AND RANDOM FOREST ALGORITHMS FOR APPLE VISION PRO SENTIMENT ANALYSIS. *Jurnal Teknik Informatika (Jutif)*, 6(1), 31–40.
<https://doi.org/10.52436/1.jutif.2025.6.1.4035>
- Riki Roska Rismansyah, Aso Sudiarjo, & Teuku Mufizar. (2025). ANALISIS SENTIMEN ULASAN SHOPEE PADA GOOGLE PLAY STORE MENGGUNAKAN ALGORITMA NAIVE BAYES. 5.
- Satya Abdul Halim Bahtiar, Chandra Kusuma Dewa, & Ahmad Luthfi. (2023). Comparison of Naïve Bayes and Logistic Regression in Sentiment Analysis on Marketplace Reviews Using Rating-Based Labeling. *Journal of Information Systems and Informatics*, 5(3), 915–927.
<https://doi.org/10.51519/journalisi.v5i3.539>
- Sukirman, Sajiah, Nursuci Putri Husain, Anastasya Febriana Syam, & Ragil Mustikosari. (2024). Analisis Sentimen Ulasan Pengguna Tiktok pada Google Play Store Berbasis TF-IDF dan Support Vector Machine. Dalam *Journal of System and Computer Engineering (JSCE) ISSN* (Vol. 5, Nomor 1).
<https://images.app.goo.gl/hC6494uW637VmYVW9>
- Tarwoto, Rizki Nugroho, Najmul Azka, & Wakhid Sayudha Rendra Graha. (2025). Analisis Sentimen Ulasan Aplikasi Mobile JKN di Google PlayStore Menggunakan IndoBERT (Vol. 3, Nomor 2).