



Analisis Sentimen pada Ulasan Aplikasi JakLingko Menggunakan Metode Naïve Bayes

Ricardus Mba Dala Pati^{1*}, Eka Kusuma Pratama², Tuslaela³

¹⁻³Universitas Bina Sarana Informatika, Indonesia

*Penulis korespondensi: rikarpaty02@gmail.com

Abstract. *JakLingko is a digital-based public transportation integration system developed to facilitate access to various transportation modes in Jakarta. Along with the increasing number of users, reviews on the JakLingko application reflect user experiences and perceptions. This study aims to analyze the sentiment of user reviews on the Google Play Store using the Naïve Bayes method. Data collection was conducted through web scraping, resulting in 3,260 reviews. The data were preprocessed, sentiment-labeled, and classified using Orange Data Mining. The research applied a quantitative experimental approach with a machine learning framework. The classification results showed that neutral sentiment dominated user reviews, followed by negative and positive sentiments. The Naïve Bayes model achieved 100% accuracy based on the confusion matrix and other evaluation metrics such as precision, recall, and F1-score. The findings highlight that Naïve Bayes can be a reliable approach for analyzing public opinion and serve as a reference for evaluating and improving digital service applications.*

Keywords: *JakLingko; Naïve Bayes; Orange; Sentiment Analysis; User Reviews.*

Abstrak. JakLingko adalah aplikasi integrasi transportasi publik berbasis digital yang dirancang untuk mempermudah akses masyarakat terhadap berbagai moda transportasi di Jakarta. Dengan bertambahnya jumlah pengguna, semakin banyak pula ulasan yang menggambarkan pengalaman dan persepsi mereka. Penelitian ini berfokus pada analisis sentimen terhadap ulasan pengguna di Google Play Store menggunakan algoritma *Naïve Bayes*. Data sebanyak 3.260 ulasan diperoleh melalui web scraping, kemudian diproses melalui tahapan preprocessing, pelabelan sentimen, serta klasifikasi menggunakan perangkat Orange Data Mining. Pendekatan yang digunakan bersifat kuantitatif eksperimental dengan basis machine learning. Hasil analisis menunjukkan bahwa sentimen netral muncul paling banyak, diikuti oleh negatif dan positif. Model *Naïve Bayes* menunjukkan performa sangat baik dengan akurasi 100% menurut confusion matrix serta metrik evaluasi lainnya, termasuk *precision*, *recall*, dan *F1-score*. Hasil ini membuktikan bahwa *Naïve Bayes* efektif untuk menganalisis opini publik berbasis teks dan dapat menjadi rujukan dalam perbaikan maupun pengembangan layanan digital.

Kata kunci: Analisis Sentimen; JakLingko; *Naïve Bayes*; Orange; Ulasan Pengguna.

1. LATAR BELAKANG

Kemajuan teknologi digital di sektor transportasi publik telah menghasilkan berbagai inovasi yang memudahkan masyarakat dalam aktivitas mobilitas sehari-hari. Salah satu contohnya adalah aplikasi JakLingko, yaitu sistem integrasi transportasi di Jakarta yang menghubungkan layanan TransJakarta, MRT, LRT, KRL, hingga Railink dalam satu platform. Kehadiran aplikasi ini bertujuan meningkatkan efisiensi perjalanan, kenyamanan, serta keterjangkauan transportasi bagi penduduk kota.

Pertumbuhan jumlah pengguna aplikasi ini memunculkan beragam ulasan dan komentar yang dipublikasikan melalui media digital, khususnya Google Play Store. Ulasan tersebut mencerminkan pengalaman nyata masyarakat, baik berupa apresiasi maupun kritik. Banyaknya jumlah komentar membuat proses penilaian secara manual menjadi tidak praktis.

Oleh karena itu, dibutuhkan metode analisis otomatis untuk memahami opini publik, salah satunya dengan pendekatan analisis sentimen.

Analisis sentimen memungkinkan pengelompokan opini ke dalam kategori positif, netral, maupun negatif. Dengan memanfaatkan algoritma Naïve Bayes, proses klasifikasi teks dapat dilakukan dengan cepat dan tetap efektif, meskipun data yang dianalisis berupa ulasan pendek. Penelitian ini menerapkan metode tersebut pada ulasan aplikasi JakLingko, dengan tujuan menggambarkan persepsi masyarakat secara lebih mendalam sekaligus memberikan masukan bagi pengembang maupun penyedia layanan transportasi digital.

2. KAJIAN TEORITIS

Analisis Sentimen

Analisis sentimen termasuk bidang kajian dalam Natural Language Processing (NLP) yang berfokus pada identifikasi opini atau sikap seseorang terhadap suatu objek. Proses ini dapat dilakukan pada level dokumen, kalimat, maupun aspek tertentu. Dalam konteks ulasan aplikasi, analisis sentimen mempermudah peneliti memahami kecenderungan opini pengguna tanpa harus menelaah komentar satu per satu. Metode ini banyak dimanfaatkan dalam penelitian mengenai kepuasan pelanggan, pemantauan reputasi suatu produk, hingga evaluasi kebijakan publik.

JakLingko

JakLingko merupakan program integrasi transportasi di Jakarta yang menyatukan berbagai moda dalam satu sistem pembayaran dan pengelolaan. Tujuannya adalah menciptakan konektivitas antarmoda yang lebih baik serta memberikan pengalaman perjalanan yang efisien dan praktis bagi masyarakat. Melalui aplikasi ini, pengguna dapat merencanakan perjalanan, membayar tiket, serta mengakses beragam transportasi hanya dengan satu platform digital. Program ini juga menjadi bagian dari strategi pemerintah daerah dalam mendukung mobilitas perkotaan yang berkelanjutan.

Metode Naïve Bayes

Naïve Bayes adalah algoritma klasifikasi berbasis probabilitas yang sering digunakan untuk pemrosesan teks. Prinsip kerjanya berlandaskan Teorema Bayes dengan asumsi independensi antar fitur. Walaupun asumsi tersebut jarang sepenuhnya terpenuhi dalam bahasa alami, metode ini tetap terbukti efektif terutama dalam mengolah data teks pendek seperti komentar media sosial maupun ulasan aplikasi. Keunggulannya meliputi kecepatan, kesederhanaan implementasi, dan performa yang stabil meskipun data latih terbatas.

Orange Data Mining

Orange adalah perangkat lunak analisis data berbasis visual yang umum dipakai dalam penelitian maupun pembelajaran. Dengan sistem alur kerja berbasis widget, pengguna dapat merancang proses analisis tanpa menulis kode. Untuk analisis teks, Orange menyediakan fitur preprocessing, ekstraksi fitur, word cloud, hingga penerapan algoritma klasifikasi seperti Naïve Bayes, k-Nearest Neighbor (kNN), dan Support Vector Machine (SVM). Antarmuka yang sederhana menjadikannya pilihan praktis dalam eksperimen machine learning terkait analisis sentimen.

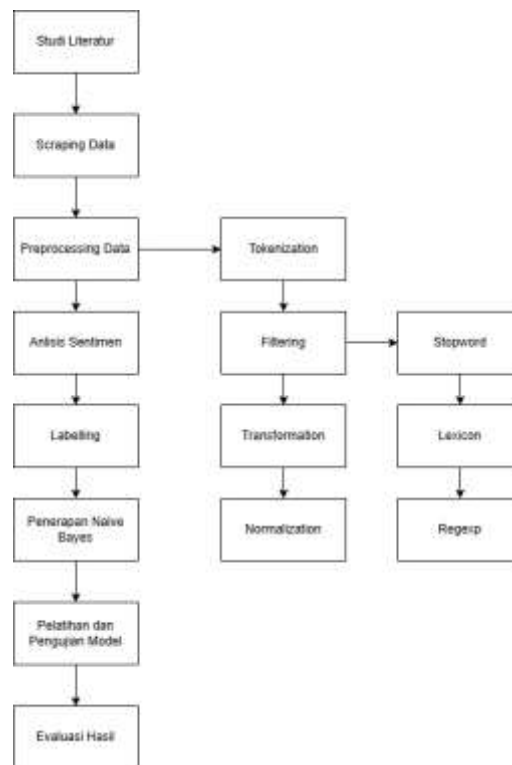
Penelitian Terkait

Sejumlah penelitian sebelumnya telah memanfaatkan Naïve Bayes untuk analisis sentimen pada aplikasi digital. Misalnya, Hasmadi et al. (2023) meneliti kualitas layanan Gojek dengan tingkat akurasi 87,4%. Nugraha & Gustian (2024) menganalisis ulasan Grab, Gojek, dan Maxim dengan akurasi tertinggi 93%. Apriliyani et al. (2024) melakukan studi pada aplikasi Duolingo dengan akurasi 86%. Sementara itu, Ladayya et al. (2022) menggunakan metode Support Vector Machine (SVM) pada data Twitter terkait JakLingko dan memperoleh akurasi 71,1%. Dari penelitian-penelitian tersebut dapat disimpulkan bahwa Naïve Bayes masih relevan digunakan dan berpotensi memberikan hasil akurasi tinggi, khususnya untuk data ulasan aplikasi di Google Play Store.

3. METODE PENELITIAN

Desain Penelitian

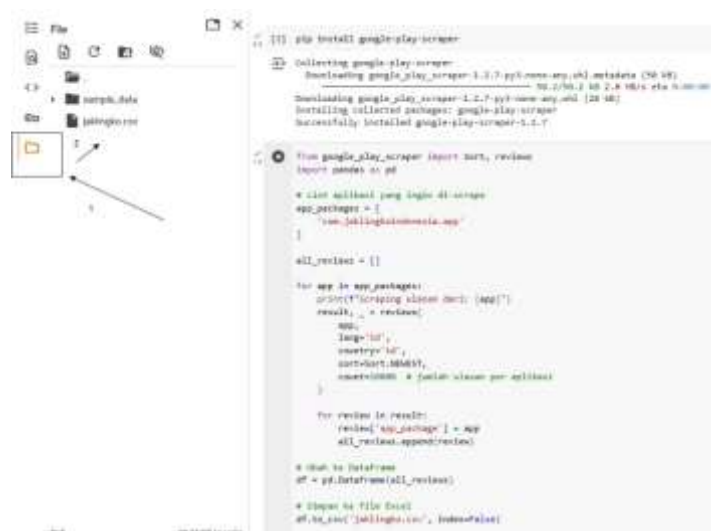
Penelitian ini menerapkan pendekatan kuantitatif dengan desain eksperimental berbasis text mining dan machine learning. Desain tersebut dipilih karena mampu mengolah data teks dalam jumlah besar serta memungkinkan pengujian algoritma klasifikasi secara terstruktur (Kowsari et al., 2019). Fokus utama penelitian adalah menganalisis sentimen ulasan pengguna aplikasi JakLingko pada Google Play Store dengan memanfaatkan algoritma Naïve Bayes.



Gambar 1. Desain Penelitian.
Sumber: Dokumen Penulis (2025).

Sumber dan Teknik Pengumpulan Data

Data penelitian diambil dari ulasan aplikasi JakLingko yang tersedia di Google Play Store. Pengumpulan data dilakukan dengan teknik web scraping menggunakan pustaka *google-play-scraper* pada Python. Dari proses ini berhasil diperoleh 3.260 ulasan berbahasa Indonesia dengan variasi opini positif, netral, dan negatif. Metode scraping dipilih karena lebih praktis dibandingkan pengumpulan manual, serta memungkinkan proses replikasi oleh peneliti lain (Hasmadi, Rahman, & Sutanto, 2023).

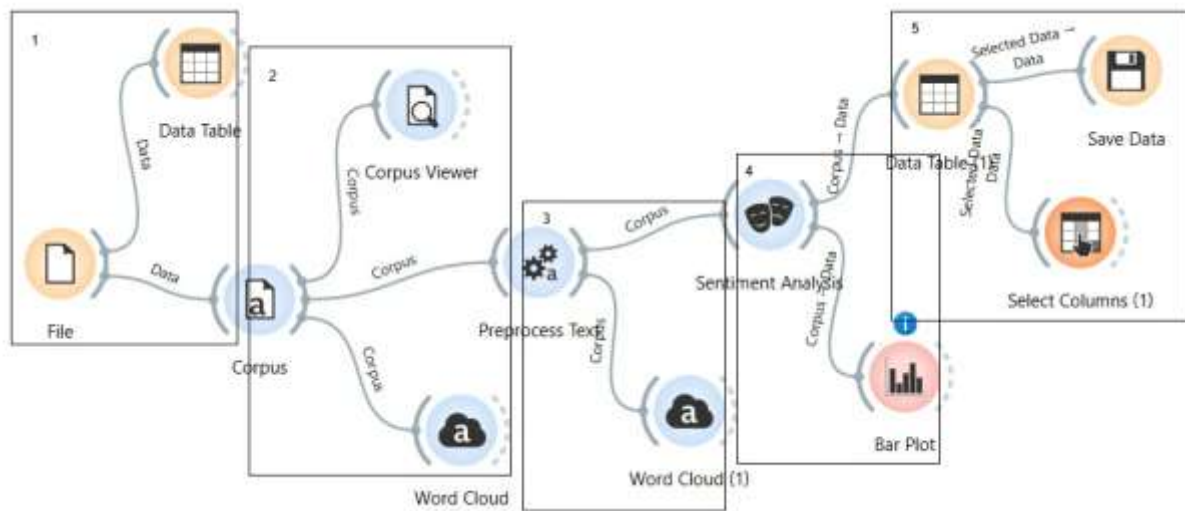


Gambar 2. Metode scraping.
Sumber: Dokumen Penulis (2025).

Tahap Preprocessing Data

Data ulasan mentah umumnya mengandung ketidakteraturan, seperti huruf kapital berlebihan, kata tidak baku, angka, maupun simbol. Untuk itu dilakukan preprocessing agar data lebih bersih dan siap dianalisis (Nugroho, Sari, & Handayani, 2020). Tahapannya meliputi: 1) Case Folding: mengubah semua huruf menjadi huruf kecil. 2) Tokenizing: memecah kalimat menjadi unit kata. 3) Stopword Removal: menghapus kata-kata umum yang tidak berkontribusi pada makna (misalnya “yang”, “dan”, “di”). 4) Stemming: mengembalikan kata ke bentuk dasarnya, misalnya “membayar”, “dibayarkan”, dan “pembayaran” menjadi “bayar”. 5) Cleaning: menghapus angka, simbol, dan karakter khusus.

Data ulasan yang diperoleh masih bersifat mentah sehingga perlu dilakukan tahap preprocessing untuk meningkatkan kualitasnya. Tahapan preprocessing mencakup case folding, tokenizing, stopwords removal, stemming, serta penghapusan angka dan tanda baca. Proses ini menghasilkan korpus teks bersih yang siap digunakan untuk analisis lebih lanjut. Tahapan preprocessing dan analisis awal dilakukan dengan bantuan perangkat lunak Orange Data Mining melalui alur kerja.



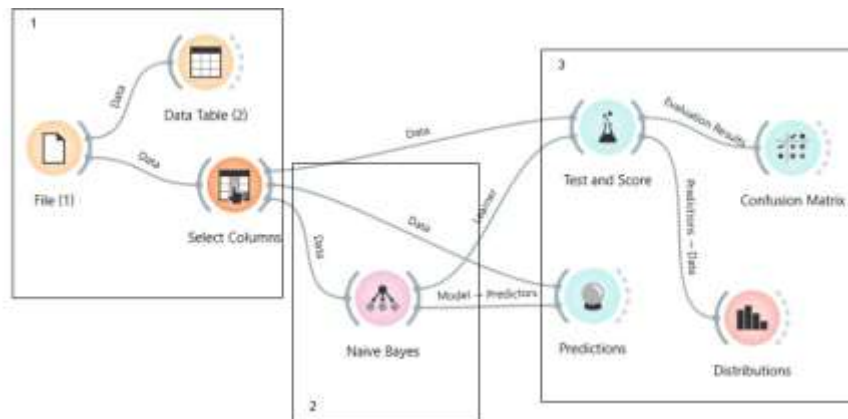
Gambar 3. Data Mining.
Sumber: Dokumen Penulis (2025).

Pelabelan Sentimen

Setiap ulasan kemudian diberi label positif, netral, atau negatif. Penentuan label dilakukan berdasarkan analisis awal polaritas kata dan konteks kalimat. Pendekatan tiga kelas ini dianggap lebih representatif dibanding klasifikasi biner karena mampu menangkap variasi opini publik (Apriliyani, Putri, & Santoso, 2024).

Algoritma Klasifikasi

Algoritma yang digunakan adalah Naïve Bayes, metode klasifikasi berbasis probabilitas yang bekerja berdasarkan Teorema Bayes dengan asumsi independensi antar fitur (Zhang & Li, 2021). Implementasi dilakukan melalui perangkat lunak Orange Data Mining yang menyediakan antarmuka visual dengan sistem *workflow*.



Gambar 4. Data Mining
Sumber: Dokumen Penulis (2025).

Implementasi algoritma dilakukan dengan Orange Data Mining, perangkat lunak berbasis visual yang mendukung workflow analisis data melalui drag-and-drop widget. Orange dipilih karena mempermudah visualisasi alur klasifikasi sekaligus menyediakan berbagai metode evaluasi performa model (Demšar et al., 2013).

Evaluasi Model

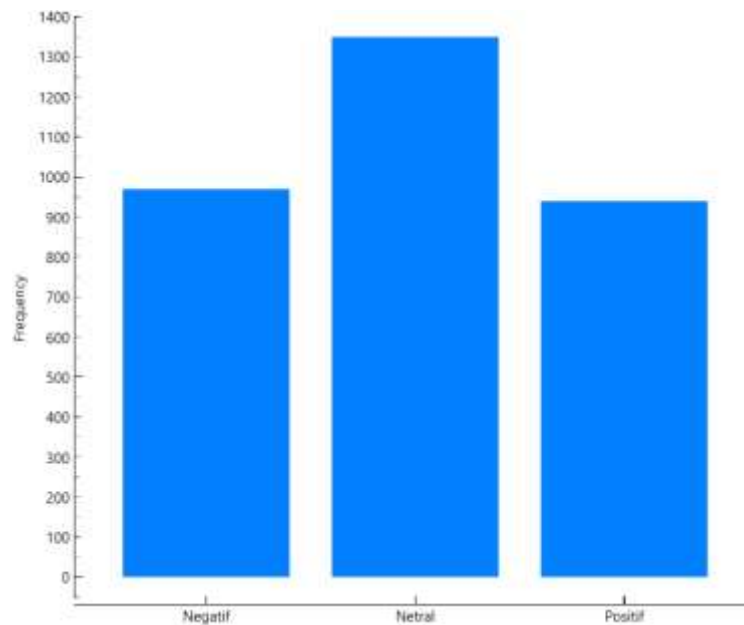
Kualitas model diukur menggunakan berbagai metrik, yaitu accuracy, precision, recall, F1-score, AUC (Area Under Curve), dan Matthews Correlation Coefficient (MCC). Evaluasi dilakukan dengan memanfaatkan confusion matrix agar distribusi hasil klasifikasi dapat dianalisis lebih detail.

Alur Penelitian

Secara garis besar, penelitian dilakukan melalui lima tahap utama: perumusan masalah, pengumpulan data ulasan, preprocessing teks, klasifikasi sentimen dengan algoritma Naïve Bayes, serta evaluasi performa model. Setelah data terkumpul, dilakukan tahap preprocessing teks yang mencakup pembersihan, normalisasi, dan penghapusan kata-kata tidak penting sehingga dihasilkan dataset yang siap digunakan. Dataset bersih tersebut kemudian diproses pada tahap klasifikasi sentimen menggunakan algoritma Naïve Bayes melalui aplikasi Orange Data Mining. Tahap terakhir adalah evaluasi performa model dan interpretasi hasil klasifikasi, di mana metrik evaluasi diterapkan untuk menilai kualitas model, sekaligus menganalisis persepsi publik berdasarkan distribusi sentimen yang dihasilkan.

4. HASIL DAN PEMBAHASAN

Dari total 3.260 ulasan pengguna aplikasi JakLingko yang berhasil dikumpulkan melalui Google Play Store, data kemudian diproses melalui tahap preprocessing dan diklasifikasikan menjadi tiga kategori sentimen: positif, netral, dan negatif. Hasil klasifikasi memperlihatkan bahwa sentimen netral paling dominan dengan jumlah 1.350 ulasan. Sementara itu, terdapat 970 ulasan bernada negatif dan 940 ulasan yang tergolong positif. Temuan ini menandakan bahwa mayoritas pengguna menuliskan komentar yang cenderung deskriptif atau bercampur antara pengalaman baik dan keluhan, sehingga tidak sepenuhnya menunjukkan sikap puas maupun kecewa.



Gambar 5. Ulasan Pengguna Aplikasi JakLingko.

Sumber: Dokumen Penulis (2025).

Proses klasifikasi dilakukan menggunakan algoritma Naïve Bayes melalui perangkat lunak Orange Data Mining. Model yang dibangun memperlihatkan performa yang sangat baik. Berdasarkan evaluasi confusion matrix, tingkat akurasi yang diperoleh mencapai 100%. Selain itu, metrik evaluasi lain seperti precision, recall, F1-score, Matthews Correlation Coefficient (MCC), dan Area Under Curve (AUC) juga memberikan hasil maksimal. Hal ini menunjukkan bahwa model tidak hanya akurat dalam mengklasifikasikan, tetapi juga konsisten dalam membedakan tiap kategori sentimen.

		Predicted			Σ
		Negatif	Netral	Positif	
Actual	Negatif	970	0	0	970
	Netral	0	1350	0	1350
	Positif	0	0	940	940
Σ		970	1350	940	3260

Gambar 6. Ulasan Aplikasi Pada Platform Google Play Store
Sumber: Dokumen Penulis (2025).

Temuan ini membuktikan bahwa algoritma Naïve Bayes efektif digunakan untuk analisis teks pendek, seperti ulasan aplikasi pada platform Google Play Store. Dibandingkan dengan penelitian sebelumnya, performa Naïve Bayes dalam studi ini bahkan lebih tinggi. Sebagai contoh, penelitian Ladayya et al. (2022) menggunakan Support Vector Machine (SVM) pada ulasan terkait JakLingko dengan akurasi 71,1%, sedangkan Nugraha & Gustian (2024) melaporkan akurasi 93% pada ulasan aplikasi transportasi daring. Oleh karena itu, penelitian ini memperkuat bukti bahwa Naïve Bayes dapat diandalkan sebagai metode analisis opini publik berbasis teks sederhana, khususnya dalam konteks layanan transportasi digital.

5. KESIMPULAN DAN SARAN

Penelitian ini mengkaji 3.260 ulasan pengguna aplikasi JakLingko dari Google Play Store dengan memanfaatkan algoritma Naïve Bayes melalui perangkat lunak Orange Data Mining. Hasil analisis menunjukkan bahwa sentimen netral mendominasi ulasan, diikuti oleh ulasan negatif dan positif. Dominasi ulasan netral menggambarkan bahwa mayoritas komentar bersifat informatif atau campuran, sehingga tidak secara jelas menunjukkan sikap puas maupun kecewa.

Algoritma Naïve Bayes terbukti memiliki kinerja yang sangat baik dengan akurasi mencapai 100% serta nilai precision, recall, F1-score, MCC, dan AUC yang juga optimal. Hal ini menegaskan bahwa Naïve Bayes efektif digunakan dalam klasifikasi teks pendek, khususnya untuk menganalisis opini pada aplikasi layanan digital.

Dari hasil penelitian ini, disarankan agar pengembang aplikasi JakLingko lebih fokus pada perbaikan aspek teknis yang kerap menjadi keluhan, seperti permasalahan login, kestabilan sistem pembayaran, dan keterbatasan fitur. Perbaikan di area tersebut diharapkan mampu mengurangi ulasan bernuansa negatif serta meningkatkan kepuasan pengguna. Sementara itu, penelitian lanjutan dapat mempertimbangkan penggunaan algoritma lain seperti Support Vector Machine, Random Forest, atau bahkan Deep Learning untuk dibandingkan performanya. Selain itu, ruang lingkup penelitian juga dapat diperluas dengan melibatkan data

dari media sosial atau forum daring agar analisis sentimen memberikan gambaran lebih menyeluruh tentang persepsi publik terhadap layanan transportasi digital.

DAFTAR REFERENSI

- Apriliyani, N., Putri, R. A., & Santoso, D. (2024). *Implementasi analisis sentimen pada ulasan aplikasi Duolingo di Google Play Store menggunakan algoritma Naïve Bayes*. *Jurnal Informatika*, 15(2), 33–42.
- Demšar, J., Curk, T., Erjavec, A., Gorup, C., Hocevar, T., Milutinovic, M., ... & Zupan, B. (2013). *Orange: Data mining toolbox in Python*. *Journal of Machine Learning Research*, 14, 2349–2353.
- Firmansyah, D. R., & Lestariningsih, E. (2024). *Analisis sentimen ulasan aplikasi SmartCampus Unisbank di Google Playstore menggunakan algoritma Naïve Bayes*. *JTIK: Jurnal Teknologi Informasi dan Komunikasi*, 8(2). Retrieved from <https://journal.lem-bagakita.org/index.php/jtik/article/view/1882>
- Hanafi, M. R., & Kurniawan, R. (2023). *Analisis sentimen pada ulasan aplikasi Sirekap di Google Play menggunakan algoritma Naïve Bayes*. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 4(4). Retrieved from <https://journal.irpi.or.id/index.php/malcom/article/view/1693>
- Hasmadi, I., Rahman, F., & Sutanto, H. (2023). *Analisis sentimen terhadap kualitas layanan driver Gojek di Play Store menggunakan algoritma Naïve Bayes dan aplikasi Orange*. *Jurnal Teknologi Informasi dan Sistem Komputer*, 11(2), 45–56.
- Heristian, S., Napiyah, M., & Erawati, W. (2025). *Analisis sentimen ulasan pelanggan menggunakan algoritma Naïve Bayes pada aplikasi Gojek*. *Computer Science (CO-SCIENCE)*, 5(1), 35–41. Retrieved from <https://jurnal.bsi.ac.id/index.php/co-science/article/view/7775>
- Kowsari, K., Meimandi, K. J., Heidarysafa, M., Mendu, S., Barnes, L., & Brown, D. (2019). *Text classification algorithms: A survey*. *Information*, 10(4), 150. <https://doi.org/10.3390/info10040150>
- Ladayya, M., Ramadhani, N., & Surya, D. (2022). *Analisis sentimen pada program transportasi publik JakLingko dengan metode Support Vector Machine*. *Jurnal Ilmu Komputer dan Informasi*, 10(3), 55–66.
- Nugraha, R., & Gustian, D. (2024). *Analisis sentimen penggunaan aplikasi transportasi online pada ulasan Google Play Store dengan metode Naïve Bayes Classifier*. *Jurnal Sistem Informasi*, 16(1), 12–25.
- Nugroho, A., Sari, D. P., & Handayani, N. (2020). *Analisis sentimen opini publik menggunakan pendekatan machine learning*. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 7(1), 55–63.
- Nurian, R., Sari, P., & Wijaya, T. (2024). *Analisis sentimen ulasan aplikasi Shopee di Google Play Store menggunakan algoritma Naïve Bayes Classifier*. *Jurnal Teknologi Informasi dan Komputer*, 12(4), 67–77.

- Putra, A., & Ramadhani, D. (2022). *Pemanfaatan analisis sentimen dalam pengukuran opini publik berbasis media sosial*. *Jurnal Sistem Informasi Indonesia*, 7(3), 112–121.
- Putri, N. A., Srirahayu, A., & Sudiby, N. A. (2024). *Sentiment analysis towards the KitaLulus application using the Naïve Bayes method from Google Play Store reviews*. *Jurnal Indonesia Sosial Teknologi*, 5(10). Retrieved from <https://jst.publikasiindonesia.id/index.php/jst/article/view/1244>
- Santoso, M. I., & Dzikrillah, A. R. (2025). *Analisis sentimen pengguna terhadap kinerja sistem transportasi umum Jakarta menggunakan algoritma Naïve Bayes*. *KLIK: Kajian Ilmiah Informatika dan Komputer*, 4(6). Retrieved from <https://djournals.com/klik/article/view/1936>
- Wijaya, E. (2021). *Machine learning untuk analisis sentimen media sosial: Studi kasus opini masyarakat*. *Jurnal Teknologi dan Sistem Informasi*, 9(2), 45–53.
- Zhang, H., & Li, X. (2021). *Naïve Bayes text classifier: A comparative study*. *Applied Artificial Intelligence*, 35(3), 205–221. <https://doi.org/10.1080/08839514.2021.1886697>