

ANALISIS SENTIMEN ULASAN APLIKASI SEABANK MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE (SVM)

Eki Ramadhan^{1*}, Ade Andri Hendriadi², Ahmad Khusaeri³

^{1,2,3}Universitas Singaperbangsa Karawang; Jl. HS. Ronggo Waluyo, Puserjaya, Telukjambe Timur, Karawang, Jawa Barat, 41361; (0267) 641177

Keywords:

Analisis Sentimen;
Seabank;
Support Vector Machine;
SMOTE;
Business Intelligence.

Correspondent Email:

ekiramadhan007@gmail.com

Abstrak. Pesatnya pertumbuhan perbankan digital di Indonesia menuntut perusahaan seperti SeaBank untuk terus mengevaluasi kualitas layanan berdasarkan persepsi pengguna. Penelitian ini bertujuan untuk menganalisis sentimen pengguna aplikasi SeaBank pada Google Play Store guna mengekstraksi wawasan Business Intelligence yang strategis. Sebanyak 71.523 ulasan mentah periode Juli hingga Desember 2025 dikumpulkan melalui teknik web scraping. Metodologi yang digunakan meliputi tahapan preprocessing teks, ekstraksi fitur TF-IDF dengan Bigram, serta penanganan data tidak seimbang menggunakan teknik Synthetic Minority Over-sampling Technique (SMOTE). Klasifikasi dilakukan menggunakan algoritma Support Vector Machine (SVM) dengan kernel Linear. Hasil pengujian menunjukkan bahwa model SVM mampu mencapai tingkat akurasi sebesar 91,21%. Temuan penelitian mengungkap dominasi sentimen positif terkait insentif bebas biaya transaksi dan kemudahan UI/UX. Sebaliknya, sentimen negatif menyoroti isu kritis pada stabilitas akses aplikasi (login) dan keterlambatan kode OTP yang berisiko memicu customer churn. Hasil analisis ini memberikan rekomendasi strategis bagi manajemen SeaBank dalam memprioritaskan perbaikan infrastruktur teknis guna meningkatkan retensi nasabah.



Copyright © [JITET](http://www.jitet.org) (Jurnal Informatika dan Teknik Elektro Terapan). This article is an open access article distributed under terms and conditions of the Creative Commons Attribution (CC BY NC)

Abstract. The rapid growth of digital banking in Indonesia requires companies like SeaBank to continuously evaluate service quality based on user perception. This study aims to analyze SeaBank application user sentiment on the Google Play Store to extract strategic Business Intelligence insights. A total of 71,523 raw reviews from July to December 2025 were collected through web scraping techniques. The methodology involves text preprocessing, TF-IDF feature extraction with Bigrams, and handling imbalanced data using the Synthetic Minority Over-sampling Technique (SMOTE). Classification was performed using the Support Vector Machine (SVM) algorithm with a Linear kernel. Test results indicate that the SVM model achieved an accuracy rate of 91.21%. The findings reveal a dominance of positive sentiment related to transaction fee incentives and UI/UX ease of use. Conversely, negative sentiment highlighted critical issues regarding application access stability (login) and OTP code delays, which risk triggering customer churn. The results of this analysis provide strategic recommendations for SeaBank management to prioritize technical infrastructure improvements to enhance customer retention.

1. PENDAHULUAN

Kemajuan teknologi informasi memicu transformasi radikal dalam industri perbankan dengan hadirnya model Bank Digital atau Neobank [1]. Sebagai salah satu bank *digital-only* di Indonesia, SeaBank sangat bergantung pada aplikasi selulernya sebagai titik kontak utama dengan nasabah [2]. Oleh karena itu, umpan balik pelanggan berupa ulasan di platform distribusi seperti Google Play Store menjadi aset data yang sangat vital. Ulasan tersebut merepresentasikan pengalaman langsung pengguna dan menjadi indikator utama kepuasan publik [3]. Namun, tingginya volume data ulasan tidak terstruktur yang dipenuhi bahasa gaul (slang), singkatan non-standar, dan typo membuat analisis secara manual menjadi tidak efisien dan rentan terhadap bias [4]. Keterlambatan dalam mengidentifikasi keluhan teknis dapat berdampak fatal pada reputasi merek dan tingkat retensi nasabah. Untuk mengatasi tantangan tersebut, pendekatan Natural Language Processing (NLP) melalui teknik text mining dan analisis sentimen sangat diperlukan sebagai instrumen Business Intelligence (BI) [5]. State of the art dari penelitian ini merujuk pada beberapa studi terdahulu yang mengimplementasikan klasifikasi sentimen pada ulasan aplikasi layanan keuangan. Penelitian terhadap aplikasi Livin' by Mandiri menggunakan algoritma Support Vector Machine (SVM) membuktikan bahwa algoritma tersebut konsisten dan sangat efektif dalam memetakan kepuasan nasabah dengan tingkat akurasi 87% [6]. Studi serupa pada aplikasi Bank Jago juga menegaskan keandalan SVM dengan akurasi rata-rata 88% [7]. Lebih lanjut, studi komparatif antara SVM dan Naïve Bayes pada dataset ulasan e-commerce dan aplikasi religi secara konsisten menunjukkan superioritas SVM dalam menangani data tekstual berdimensi tinggi tanpa asumsi independensi fitur, dengan catatan akurasi mencapai 98,93% pada studi kasus tertentu [8], [9]. Meskipun banyak studi telah dilakukan, terdapat kesenjangan (gap analysis) yang signifikan pada literatur yang ada. Mayoritas penelitian terdahulu cenderung hanya berfokus pada komparasi teknis nilai akurasi antar-algoritma, tanpa mengonversi hasil matematis

tersebut menjadi wawasan manajerial (actionable insights) yang siap dieksekusi. Selain itu, banyak studi yang mengabaikan masalah ketidakseimbangan kelas (class imbalance) ekstrem pada data natural ulasan, serta terjebak pada penggunaan ekstraksi fitur Unigram yang berisiko merusak makna kalimat berpenyangkal (negasi). Oleh karena itu, kebaruan (novelty) dari penelitian ini terletak pada pendekatan hibrida: mengintervensi ruang pelatihan menggunakan teknik Synthetic Minority Over-sampling Technique (SMOTE) untuk memastikan kepekaan sensor (Recall) terhadap kelas minoritas, menerapkan parameter Bigram untuk penyelamatan makna kontekstual negasi, serta mengadopsi kerangka kerja CRISP-DM secara utuh guna menghasilkan rekomendasi strategis Business Intelligence. Tujuan dari penelitian ini adalah merancang dan mengembangkan model klasifikasi sentimen otomatis pada ulasan aplikasi SeaBank menggunakan algoritma SVM yang dioptimasi dengan SMOTE. Selain itu, penelitian ini bertujuan untuk mengevaluasi efektivitas kinerja model secara kuantitatif melalui Confusion Matrix, serta mengidentifikasi kecenderungan topik sentimen pengguna untuk memberikan wawasan kecerdasan bisnis (BI) yang komprehensif bagi pihak manajemen SeaBank.

2. TINJAUAN PUSTAKA

2.1. Konsep Dasar Data, Informasi, dan Pengetahuan

Data adalah kumpulan fakta mentah, angka, atau simbol yang belum diolah dan belum memiliki makna kontekstual yang signifikan [2]. Ulasan pengguna pada Google Play Store merupakan representasi nyata dari data tidak terstruktur yang sarat akan opini [10]. Transformasi data mentah yang acak menjadi informasi bermakna adalah langkah krusial agar laporan memiliki relevansi kontekstual bagi penerimanya [11]. Pengetahuan (*knowledge*) merupakan tingkatan pemahaman lebih tinggi di mana informasi digabungkan dengan pengalaman untuk memahami pola dan tren guna merumuskan strategi proaktif [12].

2.2. Business Intelligence dan Data Mining

Business Intelligence (BI) memadukan arsitektur, proses, dan teknologi untuk mentransformasi data mentah menjadi pengetahuan strategis yang siap dieksekusi [13]. Penerapan BI bertujuan mengoptimalkan kualitas pengambilan keputusan bisnis agar lebih responsif dan berbasis data. *Data Mining* menempati posisi sentral dalam arsitektur BI untuk menyingkap pola laten dan mendeteksi anomali dalam himpunan data berskala besar [5]. Penggunaan algoritma statistik dan *machine learning* dalam *data mining* membantu mengekstraksi faktor-faktor yang mempengaruhi perilaku pelanggan [14].

2.3. Metodologi CRISP-DM

Cross-Industry Standard Process for Data Mining (CRISP-DM) merupakan kerangka kerja metodologis terbuka yang menjadi standar industri dalam eksekusi proyek *data science* [15]. Metodologi ini menawarkan peta jalan terstruktur yang terdiri atas enam fase fundamental: *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modeling*, *Evaluation*, serta *Deployment* [7]. Karakteristiknya yang bersifat iteratif membantu peneliti memastikan hasil analisis selaras dengan tujuan bisnis.

2.4. Text Mining, Web Scraping, dan NLP

Text Mining berfokus pada proses ekstraksi untuk menghasilkan informasi bermutu tinggi dari data teks tidak terstruktur [16], [17]. Pengumpulan data dilakukan melalui *Web Scraping*, yakni teknik ekstraksi data otomatis dari situs web ke dalam format terstruktur [18]. *Natural Language Processing* (NLP) menyediakan model teoretis untuk memahami struktur dan makna semantik dari bahasa alami manusia. Analisis sentimen merupakan penerapan spesifik NLP untuk mengukur secara kuantitatif muatan opini ke dalam kelas polaritas positif, negatif, dan netral [19].

2.5. Pra-Pemrosesan Teks dan Ekstraksi Fitur

Data teks mentah dari ulasan aplikasi seringkali "kotor" (*noisy*) sehingga memerlukan pra-pemrosesan yang meliputi *case folding*, *cleaning*, normalisasi *slang*, *stopword removal*, dan *stemming* [20]. Penerapan fitur *N-Gram* (Bigram) sangat krusial untuk menangani

fenomena negasi (*negation handling*) agar makna kalimat berpenyangkal tidak salah diklasifikasikan. Setelah bersih, teks ditransformasi menjadi representasi numerik menggunakan metode TF-IDF yang memberikan bobot berdasarkan kepentingan kata dalam korpus [21].

2.6. SMOTE dan Support Vector Machine (SVM)

Ketidakseimbangan kelas (*class imbalance*) pada data ulasan menyebabkan model cenderung mengabaikan representasi kelas minoritas [22]. Teknik *Synthetic Minority Over-sampling Technique* (SMOTE) mengatasi hal ini dengan membangkitkan sampel sintesis baru melalui interpolasi antar data terdekat di ruang fitur [23]. Algoritma SVM digunakan untuk mengidentifikasi *hyperplane* optimal guna mempartisi data dengan margin maksimum, dan terbukti efektif bekerja pada ruang berdimensi tinggi [6].

2.7. Evaluasi Model

Pasca pelatihan, evaluasi kinerja model wajib dilakukan menggunakan data uji yang belum pernah dipelajari oleh model. *Confusion Matrix* merangkum kinerja klasifikasi melalui perbandingan label aktual dengan prediksi model. Metrik akurasi, presisi, *recall*, dan *F1-score* digunakan untuk memberikan gambaran kinerja model secara menyeluruh dan objektif [9].

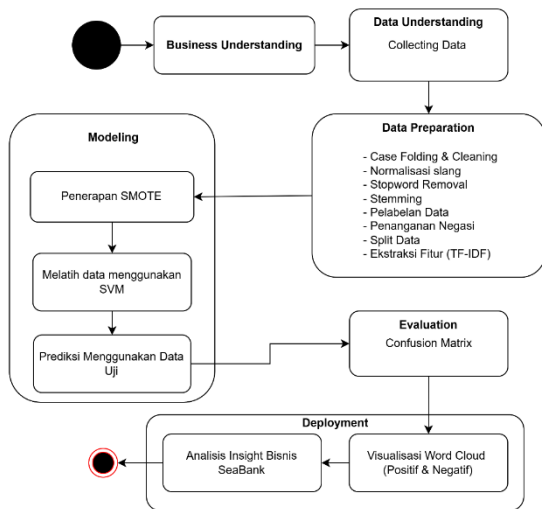
3. METODE PENELITIAN

3.1. Objek Penelitian

Objek penelitian ini adalah data tekstual tidak terstruktur berupa ulasan pengguna aplikasi perbankan digital SeaBank yang bersumber dari platform Google Play Store. Pemilihan SeaBank didasari oleh pertumbuhannya yang agresif dalam ekosistem *fintech* Indonesia, sehingga analisis ulasan menjadi krusial untuk memetakan sentimen masyarakat sebagai indikator kualitas layanan.

3.2. Rancangan Penelitian

Penelitian ini mengimplementasikan kerangka kerja standar industri *Cross-Industry Standard Process for Data Mining* (CRISP-DM).



Gambar 1. Rancangan Alur Penelitian

Metodologi ini dipilih karena pendekatannya yang terstruktur dan berorientasi pada tujuan bisnis Sistem Informasi. Tahapan penelitian dibagi menjadi enam fase utama:

1. **Business Understanding:** Mengidentifikasi kebutuhan manajemen untuk memantau sentimen ulasan secara otomatis guna menekan angka perpindahan nasabah (*customer churn*).
2. **Data Understanding:** Melakukan pengumpulan data melalui teknik *web scraping* dari Google Play Store. Total data mentah yang dikumpulkan sebanyak 71.523 ulasan periode Juli hingga Desember 2025.
3. **Data Preparation:** Melakukan pra-pemrosesan teks yang meliputi *case folding*, *cleaning*, normalisasi *slang*, *stopword removal*, dan *stemming* menggunakan pustaka Sastrawi. Pada tahap ini juga diterapkan teknik *N-Gram* (Bigram) untuk menangani kata negasi dan pelabelan otomatis berbasis rating bintang (1-2 Negatif, 4-5 Positif). Data kemudian dibagi menjadi data latih dan data uji dengan rasio 80:20.
4. **Modeling:** Mengatasi ketidakseimbangan data pada data latih menggunakan teknik *Synthetic Minority Over-sampling Technique* (SMOTE) agar proporsi kelas menjadi seimbang (50:50). Selanjutnya, dilakukan ekstraksi fitur menggunakan TF-IDF dan pelatihan model menggunakan algoritma *Support Vector Machine* (SVM) dengan kernel *Linear*.

5. **Evaluation:** Mengukur kinerja model menggunakan *Confusion Matrix* untuk mendapatkan nilai Akurasi, Presisi, *Recall*, dan *F1-Score*.

6. **Deployment:** Mentransformasi hasil klasifikasi menjadi visualisasi *Word Cloud* dan narasi rekomendasi strategis *Business Intelligence* bagi manajemen SeaBank.

4. HASIL DAN PEMBAHASAN

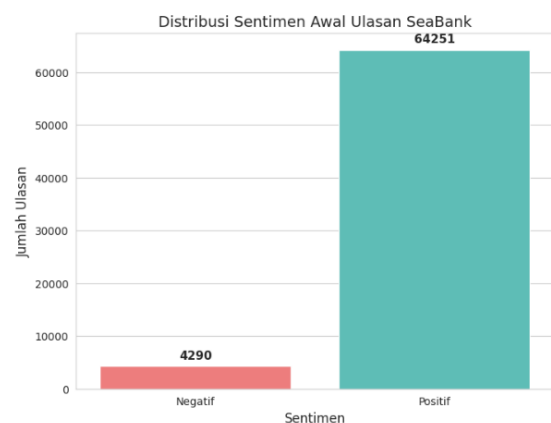
4.1. Hasil Penelitian

Hasil ekstraksi data mentah dari Google Play Store menghasilkan total 71.523 ulasan pengguna aplikasi SeaBank.

	ulasan	rating	tanggal
0	bsgus	5	2025-12-30 23:59:40
1	sangat senang karena aman , dan berbunga	5	2025-12-30 23:50:29
3	mantap gratis biaya admin	5	2025-12-30 23:38:15
4	sangat menyenangkan bagi saya	5	2025-12-30 23:33:57
5	aman dan nyaman	5	2025-12-30 23:32:50

Gambar 2. Sampel Data Mentah

Setelah melalui tahap seleksi atribut dan pembersihan ulasan kosong, diperoleh dataset bersih sebanyak 68.541 ulasan. Distribusi data awal menunjukkan dominasi sentimen positif sebanyak 64.251 ulasan dibandingkan sentimen negatif yang hanya berjumlah 4.290 ulasan.



Gambar 3. Distribusi Sentimen Awal

Proses data preparation dilakukan melalui tahapan *case folding*, *cleaning*, normalisasi *slang*, *stopword removal*, *stemming*, dan penanganan negasi menggunakan teknik Bigram. Dataset kemudian dibagi menjadi data latih dan data uji dengan rasio 80:20. Mengingat adanya ketimpangan kelas yang ekstrem, teknik SMOTE diterapkan pada data latih untuk

menyeimbangkan proporsi kelas negatif (minoritas) menjadi seimbang sempurna (50:50) sebelum dilakukan pelatihan model.

Pelatihan model menggunakan algoritma Support Vector Machine (SVM) dengan kernel Linear menunjukkan kinerja yang sangat mumpuni. Berdasarkan pengujian pada 13.709 data uji murni, model berhasil mencatatkan tingkat akurasi keseluruhan sebesar 92,25%. Hasil evaluasi mendalam menunjukkan model memperoleh tingkat precision 99% dan recall 93% untuk kelas positif, serta recall sebesar 82,40% untuk kelas negatif.

=====

HASIL EVALUASI MODEL (POS vs NEG)

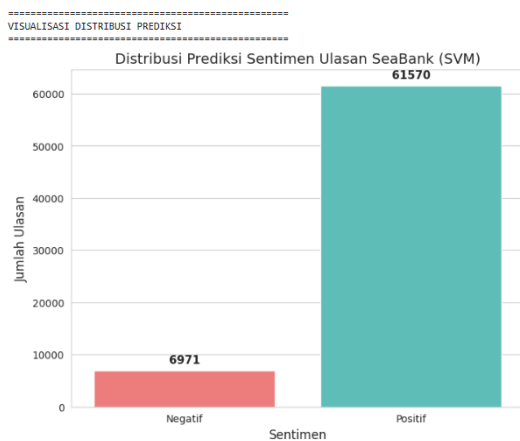
=====

Akurasi: 92.25%

	precision	recall	f1-score	support
Negatif	0.44	0.82	0.57	858
Positif	0.99	0.93	0.96	12851
accuracy			0.92	13709
macro avg	0.71	0.88	0.76	13709
weighted avg	0.95	0.92	0.93	13709

Gambar 4. Hasil Classification Report

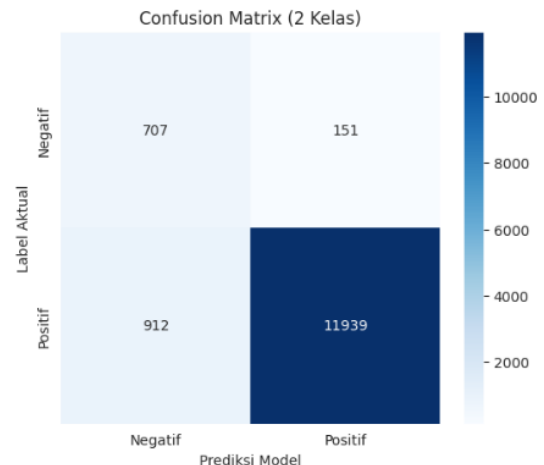
Berdasarkan hasil prediksi sentimen yang telah dilakukan secara menyeluruh terhadap 68.541 dokumen ulasan menggunakan algoritma Support Vector Machine (SVM), diperoleh gambaran komprehensif mengenai persepsi pengguna. Analisis distribusi sentimen memperlihatkan dominasi kepuasan pengguna yang sangat signifikan, di mana teridentifikasi sebanyak 61.570 ulasan bersentimen positif dibandingkan dengan 6.971 ulasan bersentimen negatif, sebagaimana direpresentasikan pada Gambar 5.



Gambar 5. Hasil Distribusi Sentimen

4.2. Pembahasan dan Wawasan Bisnis

Keberhasilan model dalam mendeteksi keluhan nasabah dipengaruhi oleh penggunaan SMOTE yang berhasil meningkatkan sensitivitas (recall) kelas negatif secara signifikan.



Gambar 6. Hasil Confusion Matrix

Meskipun terjadi fenomena precision-recall trade-off pada kelas negatif (presisi 44%), hal ini merupakan kompromi operasional yang dapat diterima dalam perspektif Business Intelligence guna meminimalkan risiko kehilangan nasabah akibat keluhan kritis yang terlewat.

Analisis wawasan bisnis melalui frekuensi kata dan visualisasi Word Cloud mengungkap determinan utama kepuasan pengguna SeaBank.



Gambar 7. Hasil Word Cloud Positif

Sentimen positif didominasi oleh kata kunci "gratis", "mudah", dan "cepat", yang membuktikan bahwa strategi insentif biaya administrasi dan kualitas desain UI/UX aplikasi telah tepat sasaran.



Gambar 8. Hasil Word Cloud Negatif

Sebaliknya, sentimen negatif mengerucut pada isu instabilitas peladen (server) yang menyebabkan kendala akses (login) dan kegagalan pengiriman kode OTP saat proses pendaftaran. Isu teknis pada gerbang pendaftaran ini membawa implikasi serius terhadap potensi customer churn dan inefisiensi biaya akuisisi pelanggan (Customer Acquisition Cost). Oleh karena itu, peningkatan kapasitas infrastruktur dan evaluasi vendor layanan SMS OTP menjadi rekomendasi prioritas bagi manajemen SeaBank.

5. KESIMPULAN

Berdasarkan hasil analisis sentimen ulasan aplikasi SeaBank, dapat ditarik beberapa kesimpulan utama:

- a. Pengembangan model klasifikasi sentimen telah berhasil diimplementasikan sesuai kerangka kerja CRISP-DM melalui tahapan *data preparation* yang ketat terhadap 68.541 ulasan bersih, termasuk normalisasi *slang* dan penanganan negasi menggunakan *bigram*.
- b. Penerapan algoritma *Support Vector Machine* (SVM) yang dipadukan dengan teknik SMOTE terbukti sangat efektif dengan mencatatkan akurasi sebesar 92,25%. Nilai *recall* kelas negatif yang mencapai 82,40% memvalidasi bahwa sistem memiliki kepekaan tinggi dalam mendeteksi keluhan nasabah dari tumpukan data mayoritas.
- c. Determinan utama sentimen positif dipicu oleh insentif layanan gratis serta kualitas desain UI/UX yang memudahkan transaksi. Sebaliknya, sentimen negatif mengerucut pada isu instabilitas *server* yang menghambat akses *login* dan kendala pengiriman kode OTP pada proses pendaftaran.

- d. Manajemen SeaBank direkomendasikan melakukan audit teknis terhadap infrastruktur peladen guna menekan angka kegagalan akses aplikasi dan mengevaluasi vendor penyedia layanan SMS OTP untuk memperlancar proses verifikasi calon nasabah. Peneliti selanjutnya disarankan untuk menguji coba algoritma berbasis Deep Learning seperti LSTM atau model pre-trained seperti IndoBERT untuk menangkap konteks semantik (sarkasme) dengan lebih akurat. Selain itu, penerapan klasifikasi multi-class (termasuk sentimen netral) dapat dilakukan untuk mendapatkan peta persaingan bisnis yang lebih komprehensif.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada pihak-pihak terkait yang telah memberi dukungan terhadap penelitian ini.

DAFTAR PUSTAKA

- [1] OJK, “Transformasi Digital Perbankan.”
- [2] F. Adelia Irawan, A. Rialdy Atmadja, and A. Wahana, “Analisis Sentimen Ulasan Aplikasi Bank Digital Menggunakan Algoritma Naive Bayes,” *Journal of Computer Science and Information Technology E-ISSN*, vol. 4, no. 2, pp. 60–68, 2024.
- [3] P. P. Wulan and H. Basri, “ANALISIS SENTIMEN TERHADAP LAYANAN NASABAH BANK MENGGUNAKAN TEKNIK KLASIFIKASI NAIVE BAYES SENTIMENT ANALYSIS OF BANKING CUSTOMER SERVICE USING NAIVE BAYES CLASSIFICATION TECHNIQUE,” *Jurnal Kecerdasan Buatan dan Teknologi Informasi*, vol. 3, no. 2, pp. 68–74, 2024.
- [4] R. C. Tarumingkeng, “Natural Language Processing (NLP),” Bogor, Nov. 2024.
- [5] S. Setiyani, A. Aminah, and E. Rilvani, “PERAN DAN STRATEGI DATA MINING DALAM TRANSFORMASI BISNIS PERUSAHAAN,” *Jurnal Media Akademik (JMA)*, vol. 3, no. 8, Aug. 2025, doi: 10.62281.
- [6] 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.
- [7] L. Anroza Putra, “ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI BANK

- JAGO MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE,” UNIVERSITAS ISLAM NEGERI SULTAN SYARIF KASIM RIAU, 2024.
- [8] A. M. Rizki, B. Bustami, and S. F. Anshari, “Comparison of Support Vector Machine and Naïve Bayes Algorithms in Sentiment Analysis of Tiktokshop Application User Reviews,” *Journal of Renewable Energy, Electrical, and Computer Engineering*, vol. 5, no. 1, pp. 18–29, Mar. 2025, doi: 10.29103/jreece.v5i1.21342.
- [9] H. Aprilianti, H. Mustofa, K. Umam, and M. R. Handayani, “Comparative Study of SVM, KNN, and Naïve Bayes for Sentiment Analysis of Religious Application Reviews,” *Journal of Applied Informatics and Computing (JAIC)*, vol. 9, no. 3, Jun. 2025, [Online]. Available: <http://jurnal.polibatam.ac.id/index.php/JAIC>
- [10] A. Hilman Hakim *et al.*, “Analisis Sentimen Aplikasi Bank Digital Pada Google Play Store Menggunakan Algoritma Naive Bayes,” *JUNI 2025 Journal of Informatics and Communications Technology*, vol. 7, no. 1, pp. 23–034, doi: 10.52661.
- [11] Maydianto and M. Rasid Ridho, “RANCANG BANGUN SISTEM INFORMASI POINT OF SALE DENGAN FRAMEWORK CODEIGNITER PADA CV POWERSHOP,” *JURNAL COMASIE*, vol. 4, 2021.
- [12] S. Ikhlasia Kasim, R. Hafid, and R. Wahyuni Mohamad, “Hubungan Tingkat Pengetahuan tentang Kesehatan Reproduksi dengan Kesiapan Menghadapi Masa Pubertas Pada Remaja Usia 12-13 Tahun Di SMP Negeri 1 Limboto,” *Jurnal Kolaboratif Sains*, vol. 8, no. 4, pp. 1769–1784, Apr. 2025, doi: 10.56338/jks.v8i4.7177.
- [13] D. Rizki and A. I. Komputer, “TEXT MINING UNTUK PENGAMBILAN KEPUTUSAN DALAM RISET PASAR,” 2024.
- [14] H. Ar-Rasyid, A. Fitriansyah, J. Sari, and I. Ramadhan, “Sistem Rekomendasi Penjualan Cross-Selling Potensial Dengan Pendekatan Data Mining.”
- [15] R. R. Aria, “Implementasi Algoritma K-Means untuk Pengelompokan Data Imunisasi Balita dengan Metode CRISP-DM,” *Remik: Riset dan E-Jurnal Manajemen Informatika Komputer*, vol. 9, no. 1, pp. 189–197, Jan. 2025, doi: 10.33395/remik.v9i1.14391.
- [16] A. Nur Syamsudin and U. Budiyanto, “Analisis Sentimen Saran Pengguna Mandatory E-Learning Menggunakan Text Mining pada Learning Management System,” *Technomedia Journal*, vol. 9, no. 3, pp. 346–359, Feb. 2025, doi: 10.33050/tmj.v9i3.2368.
- [17] G. Rafianto, Hannie, and A. Ma’sum, “ANALISIS SENTIMEN BERBASIS TOPIK ULASAN PENGGUNA APLIKASI ROBLOX MENGGUNAKAN INTEGRASI LDA DAN SVM,” *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 14, no. 2, Apr. 2026, doi: 10.23960/jitet.v14i2.9336.
- [18] R. Hasanah, Z. Hasanah, S. Ardi Wijaya, D. Abdennasser, and A.-N. Sharkawy, “Play Store Data Scrapping and Preprocessing done as Sentiment Analysis Material,” 2025. [Online]. Available: <https://journal.abhinaya.co.id/index.php/IJMS>
- [19] P. P. Wulan and H. Basri, “ANALISIS SENTIMEN TERHADAP LAYANAN NASABAH BANK MENGGUNAKAN TEKNIK KLASIFIKASI NAIVE BAYES,” *Jurnal Kecerdasan Buatan dan Teknologi Informasi*, vol. 3, no. 2, pp. 68–74, May 2024.
- [20] H. Wijaya and N. Hayati, “NATURAL LANGUAGE PROCESSING (NLP) UNTUK ANALISIS SENTIMEN ULASAN SEBLAK BANDUNG PEDAS KUDUS,” *Jurnal of Business and Audit Information System (JBASE)*, vol. 8, no. 1, pp. 13–22, 2025, doi: 10.30813/jbase.v8i1.8035.
- [21] F. Amandasari and D. Damayanti, “Perbandingan Kinerja Support Vector Machine dan Naive Bayes dalam Klasifikasi Sentimen Twitter Terhadap Pelayanan BPJS,” *Jurnal Pendidikan dan Teknologi Indonesia*, vol. 5, no. 3, pp. 645–653, Mar. 2025, doi: 10.52436/1.jpti.680.
- [22] I. de Zarzà, J. de Curtò, and C. T. Calafate, “Optimizing Neural Networks for Imbalanced Data,” *Electronics (Switzerland)*, vol. 12, no. 12, Jun. 2023, doi: 10.3390/electronics12122674.
- [23] S. DEMİR and E. K. ŞAHİN, “Evaluation of Oversampling Methods (OVER, SMOTE, and ROSE) in Classifying Soil Liquefaction Dataset based on SVM, RF, and Naïve Bayes,” *European Journal of Science and Technology*, Feb. 2022, doi: 10.31590/ejosat.1077867.