



ANALISIS SENTIMEN PENGGUNA APLIKASI LAYANAN PUBLIK BERBASIS WEB MENGGUNAKAN KOMPARASI ALGORITMA SUPPORT VECTOR MACHINE DAN NAÏVE BAYES

Patrick Wellman Simbolon, Andhi Mahendra Ch. Sumarjo, Feike Mecin Lamia

Universitas Parna Raya, Manado

patrick@microsoft.parnaraya.ac.id

andhisumarho@microsoft.parnaraya.ac.id

feikelamia@microsoft.parnaraya.ac.id

Abstrak

Transformasi digital pada sektor pemerintahan telah mendorong pemanfaatan aplikasi layanan publik berbasis web guna meningkatkan efisiensi dan transparansi birokrasi. Namun demikian, keberhasilan implementasi platform digital tersebut sangat bergantung pada persepsi, kenyamanan, dan kepuasan masyarakat selaku pengguna akhir. Penelitian ini bertujuan untuk menganalisis sentimen masyarakat terhadap kualitas layanan aplikasi web pelayanan publik dan membandingkan performa dua algoritma klasifikasi teks populer, yaitu Support Vector Machine (SVM) dan Multinomial Naïve Bayes. Data ulasan sebanyak 1.000 entri dikumpulkan dan diproses melalui tahapan text preprocessing yang komprehensif, mencakup case folding, data cleansing, tokenizing, stopword removal berbasis Sastrawi, dan stemming dengan algoritma Nazief-Adriani. Ekstraksi fitur dilakukan menggunakan metode Term Frequency-Inverse Document Frequency (TF-IDF) dengan rentang unigram dan bigram. Pengujian model dievaluasi secara ketat menggunakan metode 10-fold cross-validation untuk menjamin validitas dan stabilitas hasil komputasi. Hasil eksperimen menunjukkan bahwa algoritma Support Vector Machine dengan kernel linear mengungguli Multinomial Naïve Bayes pada seluruh metrik pengujian. SVM memperoleh tingkat akurasi sebesar 99,80%, presisi 100,00%, recall 99,60%, dan F1-score sebesar 99,80%. Sementara itu, Multinomial Naïve Bayes mencatatkan akurasi sebesar 94,00%, presisi 93,90%, recall 94,20%, dan F1-score sebesar 94,00%. Analisis matriks konfusi membuktikan ketangguhan SVM dalam memisahkan bidang batas hyperplane pada fitur berdimensi tinggi, serta mampu meminimalkan kesalahan klasifikasi pada opini berkalimat kompleks. Temuan penelitian ini memberikan kontribusi teoretis dalam pemodelan NLP bahasa Indonesia serta memberikan wawasan strategis bagi pengelola website layanan publik dalam memprioritaskan perbaikan infrastruktur server, responsivitas antarmuka, dan transparansi pelacakan permohonan berkas.

Kata Kunci: Analisis Sentimen, Layanan Publik Web, Support Vector Machine, Naïve Bayes, TF-IDF, Cross Validation

Abstract

The digital transformation in the public sector has driven the widespread adoption of web-based public service applications to enhance bureaucratic efficiency and transparency. Nevertheless, the ultimate success of these digital platforms relies heavily on the perceptions, comfort, and satisfaction of citizens as end-users. This study aims to analyze public sentiment toward web-based public service applications and compare the performance of two prominent text classification algorithms: Support Vector Machine (SVM) and Multinomial Naïve Bayes. A dataset comprising 1,000 user reviews was gathered and processed through comprehensive natural language preprocessing stages, including case folding, cleansing, tokenizing, Sastrawi-based stopword removal, and Nazief-Adriani stemming. Feature extraction was implemented using Term Frequency-Inverse Document Frequency (TF-IDF) spanning unigram and bigram ranges. Model validation was systematically conducted via 10-fold cross-validation to guarantee empirical reliability. Experimental findings demonstrate that the linear Support Vector Machine consistently surpassed Multinomial Naïve Bayes across all evaluated metrics. SVM achieved an accuracy of 99.80%, precision of 100.00%, recall of 99.60%, and an F1-score of 99.80%. In comparison, Multinomial Naïve Bayes achieved an accuracy of 94.00%, precision of 93.90%, recall of 94.20%, and an F1-score of 94.00%. Confusion matrix analysis confirms SVM's superior capability in constructing optimal separating hyperplanes within high-dimensional sparse text vectors, thereby mitigating misclassification on syntactically complex opinions. These insights provide both theoretical contributions to Indonesian NLP and pragmatic guidelines for public web administrators to refine server capacity, user interface responsiveness, and administrative tracking transparency.

Keywords: Sentiment Analysis, Web Public Services, Support Vector Machine, Naïve Bayes, TF-IDF, Cross Validation



I. PENDAHULUAN

Penerapan sistem pemerintahan berbasis elektronik (e-government) di era digital telah menjadi instrumen fundamental dalam modernisasi tata kelola administrasi publik di Indonesia (Santoso & Nugroho, 2021). Transformasi ini diwujudkan melalui pengembangan berbagai aplikasi pelayanan publik berbasis web yang dirancang untuk mempermudah masyarakat dalam mengurus dokumen kependudukan, perizinan berusaha, sertifikasi, hingga administrasi ketenagakerjaan secara mandiri (Kurniawan & Utami, 2022). Pemanfaatan sistem berbasis web menawarkan keunggulan berupa kemudahan aksesibilitas tanpa terbatas ruang dan waktu, pemangkasan birokrasi berbelit, serta pengurangan potensi pungutan liar melalui sistem verifikasi digital yang terintegrasi (Aktama et al., 2025; Akriyono et al., 2025).

Meskipun demikian, peralihan dari pelayanan konvensional menuju sistem digital berbasis web tidak terlepas dari sejumlah kendala teknis dan operasional. Dalam implementasi praktisnya di lapangan, pengguna kerap mengeluhkan performa server yang lambat (loading time tinggi), seringnya terjadi galat (error) pada saat pengunggahan berkas persyaratan, keterlambatan penerimaan kode verifikasi satu kali (One-Time Password/OTP), hingga desain antarmuka pengguna (User Interface/UI) yang kurang intuitif bagi masyarakat awam (Fahrezi et al., 2022). Keluhan, masukan, dan pujian dari masyarakat umumnya disampaikan secara tertulis melalui portal evaluasi, formulir umpan balik, media sosial resmi, maupun forum diskusi daring. Timbunan teks opini publik tersebut merepresentasikan data berharga yang sangat penting untuk diekstraksi guna mengetahui tingkat kepuasan masyarakat secara objektif (Wahyuningrum, 2021).

Untuk mengolah data opini masyarakat dalam volume besar, metode pembacaan dan rekapitulasi manual terbukti tidak efisien, membutuhkan waktu lama, serta rentan terhadap bias subjektivitas evaluator. Oleh karena itu, pendekatan Pemrosesan Bahasa Alami (Natural Language Processing/NLP) melalui teknik Analisis Sentimen (Sentiment Analysis) atau penambangan opini (opinion mining) menjadi solusi komputasional yang sangat relevan (Cahyono & Suharyanto, 2022). Analisis sentimen memungkinkan sistem komputer mengidentifikasi, mengekstraksi, dan mengklasifikasikan polaritas polaritas opini publik ke dalam kategori positif atau negatif secara otomatis dan akurat.

Dalam domain klasifikasi teks, algoritma Support Vector Machine (SVM) dan Naïve Bayes merupakan dua metode pembelajaran mesin (machine learning) berbasis pengawasan (supervised learning) yang paling populer dan sering diandalkan (Pratama et al., 2023). Algoritma Support Vector Machine dikenal tangguh dalam menangani data berdimensi tinggi seperti representasi teks, memiliki kemampuan generalisasi yang sangat baik, serta mampu mencegah terjadinya overfitting dengan memaksimalkan margin bidang pemisah (separating hyperplane) (Setiawan et al., 2021). Di sisi lain, Multinomial Naïve Bayes merupakan pengklasifikasi probabilistik berbasis Teorema Bayes yang memiliki keunggulan dalam kecepatan komputasi, kesederhanaan matematis, dan efisiensi memori yang tinggi (Wahyudi et al., 2021).

Meskipun kedua algoritma tersebut telah diterapkan pada berbagai kasus analisis sentimen, perbandingan komparatif keduanya secara mendalam pada domain ulasan aplikasi web pelayanan publik berbahasa Indonesia masih memerlukan eksplorasi empiris yang lebih terperinci (Rahmad et al., 2022). Karakteristik teks bahasa Indonesia pada ulasan layanan publik seringkali memuat sinonimitas yang luas, penggunaan partikel penegas, hingga kalimat majemuk bertingkat yang memadukan kritik teknis dengan apresiasi layanan (Iriananda & Rangkuti, 2021). Oleh sebab itu, penelitian ini bertujuan untuk mengevaluasi dan membandingkan performa model klasifikasi Support Vector Machine dan Multinomial Naïve Bayes dalam mengklasifikasikan sentimen pengguna aplikasi web layanan publik. Hasil penelitian ini diharapkan tidak hanya memberikan kontribusi teoritis bagi literatur penambangan



teks bahasa Indonesia, tetapi juga menjadi dasar pertimbangan strategis bagi pengembang perangkat lunak pemerintah dan pengambil kebijakan dalam meningkatkan kualitas layanan publik berbasis web.

II. METODE PENELITIAN

Penelitian ini dilaksanakan melalui pendekatan kuantitatif eksperimental yang terbagi ke dalam enam tahapan utama: (1) Pengumpulan Dataset Ulasan, (2) Text Preprocessing, (3) Pembobotan Kata Term Frequency-Inverse Document Frequency (TF-IDF), (4) Pelatihan Model Klasifikasi SVM dan Naïve Bayes, (5) Evaluasi Model Menggunakan 10-Fold Cross-Validation, dan (6) Analisis Perbandingan Performa Komputasi.

2.1 Pengumpulan Dataset Ulasan

Dataset yang digunakan dalam penelitian ini terdiri dari 1.000 ulasan pengguna terhadap aplikasi pelayanan publik berbasis web di Indonesia. Data tersebut mencakup opini masyarakat terkait keandalan sistem, kecepatan respon, kemudahan navigasi antarmuka, serta efektivitas pelayanan berkas. Dataset memiliki komposisi seimbang (balanced dataset) yang terdiri dari 500 ulasan berlabel positif (skor sentimen 1) dan 500 ulasan berlabel negatif (skor sentimen 0) guna mencegah terjadinya bias mayoritas kelas (class imbalance problem) pada tahap pelatihan model pembelajaran mesin (Anggraini et al., 2020).

2.2 Text Preprocessing

Teks ulasan mentah dari pengguna umumnya memuat derau (noise), tanda baca, huruf kapital yang tidak konsisten, serta kata-kata penghubung yang tidak memiliki muatan informasi sentimen. Oleh karena itu, tahapan text preprocessing diterapkan secara sistematis melalui urutan berikut:

1. Case Folding: Mengubah seluruh karakter huruf dalam teks menjadi huruf kecil (lowercase) untuk memastikan keseragaman representasi karakter.

2. Data Cleansing: Membersihkan teks dari elemen-elemen derau, seperti tanda baca, angka, simbol matematika, pranala (URL), tanda tagar, serta spasi berlebih (whitespace trimming).

3. Tokenizing: Memecah susunan kalimat menjadi unit-unit kata individual (token) guna memfasilitasi analisis leksikal lebih lanjut.

4. Stopword Removal: Menghilangkan kata-kata umum yang memiliki frekuensi kemunculan tinggi namun tidak memiliki bobot sentimen signifikan (contoh: 'yang', 'dan', 'di', 'dari', 'ke', 'ini', 'itu'). Proses ini mengacu pada kamus stopwords Sastrawi untuk teks bahasa Indonesia.

5. Stemming: Mengubah kata berimbuhan menjadi bentuk kata dasarnya (root word) menggunakan algoritma Nazief-Adriani yang dirancang khusus untuk morfologi bahasa Indonesia yang kaya imbuhan (prefiks, sufiks, infiks, dan konfiks).

2.3 Ekstraksi Fitur TF-IDF (Term Frequency-Inverse Document Frequency)

Setelah tahap pembersihan teks selesai, data teks harus dikonversi ke dalam vektor representasi numerik. Metode yang digunakan adalah TF-IDF, yang mengukur signifikansi suatu kata t dalam suatu dokumen d relatif terhadap keseluruhan korpus dokumen D (Putra & Rusjyanthi, 2021; Suryani et al., 2022). Nilai TF-IDF dihitung melalui perkalian Term Frequency (TF) dan Inverse Document Frequency (IDF) sebagaimana dirumuskan pada persamaan (1) dan (2):

$$IDF(t, D) = \log (|D| / (1 + |\{d \text{ in } D : t \text{ in } d\}|)) \quad (1)$$

$$W(t, d) = TF(t, d) * IDF(t, D) \quad (2)$$

Di mana $TF(t, d)$ adalah frekuensi kemunculan kata t dalam dokumen d , $|D|$ merepresentasikan total dokumen dalam korpus (1.000 ulasan), dan penyebut merepresentasikan jumlah dokumen yang memuat kata t ditambah konstanta penghalusan Laplace.

2.4 Algoritma Support Vector Machine (SVM)

Support Vector Machine merupakan metode supervised learning yang bertujuan mencari bidang pembatas pemisah (optimal separating hyperplane) yang memisahkan dua kelas data dengan jarak margin terbesar (Marlina & Hasan, 2022). Diberikan data pelatihan (x_i, y_i) di mana x_i merupakan vektor fitur TF-IDF dan y_i in $\{-1, +1\}$ adalah label kelas sentimen. Persamaan hyperplane pemisah didefinisikan sebagai:

$$w \cdot x + b = 0 \quad (3)$$

Di mana w adalah vektor bobot berarah tegak lurus terhadap bidang pembatas dan b adalah nilai bias. Optimasi fungsi tujuan SVM dengan parameter penalti C dirumuskan melalui pemrograman kuadratik:

$$\min 0.5 * ||w||^2 + C * \sum(x_i) \quad (4)$$

Dengan konstrain kendala: $y_i(w \cdot x_i + b) \geq 1 - x_i$, di mana $x_i \geq 0$ merupakan variabel slack untuk mengakomodasi titik data yang berada di dalam margin.

2.5 Algoritma Multinomial Naïve Bayes

Multinomial Naïve Bayes merupakan pengklasifikasi probabilistik yang mengasumsikan independensi kondisional antar fitur teks (asumsi naif). Probabilitas posterior dokumen d terhadap kelas sentimen c in $\{\text{positif, negatif}\}$ dihitung berdasarkan Teorema Bayes:

$$P(c|d) = [P(c) * \prod(P(t_k|c))] / P(d) \quad (5)$$

Estimasi probabilitas kemunculan fitur kata t_k pada kelas c dihitung menggunakan teknik penghalusan Laplace (Laplace smoothing) dengan parameter $\alpha = 1.0$:

$$P(t_k|c) = (N_{ck} + \alpha) / (N_c + \alpha * |V|) \quad (6)$$

Di mana N_{ck} adalah total frekuensi kemunculan kata t_k pada dokumen kelas c , N_c adalah jumlah seluruh kata dalam kelas c , dan $|V|$ adalah ukuran perbendaharaan kosakata unik korpus.

2.6 Skema Evaluasi dan Validasi Silang (10-Fold Cross-Validation)

Untuk menghindari bias partisi data pada pemodelan, pengujian dilakukan dengan skema 10-Fold Cross-Validation (Merawati et al., 2022). Dataset 1.000 ulasan dibagi secara acak ke dalam 10 subset berukuran sama (100 ulasan per fold). Pada setiap iterasi, 9 fold dialokasikan sebagai data latih (training set) dan 1 fold dialokasikan sebagai data uji (testing set). Proses diulang sebanyak 10 kali sehingga setiap fold pernah menjadi data uji. Metrik performa yang dievaluasi mencakup Akurasi (Accuracy), Presisi (Precision), Sensitivitas (Recall), dan Nilai F1 (F1-Score) berbasis Confusion Matrix.

III. HASIL DAN PEMBAHASAN

Bagian ini memaparkan hasil pengujian komputasi yang telah dilakukan terhadap 1.000 ulasan aplikasi layanan publik berbasis web, meliputi analisis leksikal fitur TF-IDF, hasil komparasi performa algoritma, evaluasi matriks kebingungan, serta analisis implikasi praktis.

3.1 Analisis Fitur Kata Bobot Sentimen Teratas (Top TF-IDF Features)

Berdasarkan proses pelatihan model linear Support Vector Machine, setiap fitur kata memperoleh nilai koefisien bobot yang mengindikasikan tingkat signifikansinya terhadap polaritas sentimen positif maupun negatif. Fitur kata dengan bobot positif tertinggi merepresentasikan indikator kepuasan pengguna, sedangkan bobot negatif tertinggi mencerminkan aspek-aspek yang menimbulkan ketidakpuasan masyarakat.

Tabel 1. Sepuluh Fitur Kata dengan Bobot Koefisien Terbesar

No	Fitur Positif (Bobot)	Fitur Negatif (Bobot)	Dimensi Layanan Web Terkait
1	layanan cepat (+1.77)	gangguan sistem (-1.73)	Stabilitas & Waktu Respon Server
2	mudah dipahami (+1.26)	membingungkan (-1.69)	Desain Navigasi & Usability UI
3	sangat terbantu (+1.01)	sering lambat (-1.63)	Efektivitas Pemrosesan Berkas
4	verifikasi kilat (+0.99)	gagal unggah (-1.39)	Keandalan Modul Lampiran File
5	antarmuka bersih (+0.96)	pranala rusak (-1.28)	Kerapian Tata Letak & Link
6	menghemat waktu (+0.92)	formulir reset (-1.04)	Penyimpanan Sesi Pengguna
7	penerbitan praktis (+0.89)	otp terlambat (-0.90)	Keandalan Gateway SMS/Email
8	fitur tracking (+0.84)	data tidak sinkron (-0.85)	Integrasi Basis Data Pusat
9	admin tanggap (+0.81)	tetap berkas fisik (-0.75)	Integrasi Layanan End-to-End
10	hemat biaya (+0.78)	batas file kecil (-0.75)	Konfigurasi Kuota Unggah Dokumen

Dari Tabel 1, terlihat bahwa aspek kepuasan masyarakat terhadap aplikasi web sangat didorong oleh indikator kecepatan verifikasi, penghematan waktu, dan kejelasan navigasi antarmuka. Sebaliknya, pemicu utama ketidakpuasan masyarakat berpusat pada kegagalan teknis saat mengunggah berkas, keterlambatan penerimaan kode OTP, dan ketidaksinkronan data kependudukan yang mengharuskan pemohon tetap datang membawa dokumen fisik ke kantor pelayanan.

3.2 Perbandingan Kinerja Algoritma SVM vs Naïve Bayes

Evaluasi kinerja kedua model pembelajaran mesin dilakukan berdasarkan skema 10-fold cross validation. Rata-rata metrik evaluasi yang diperoleh pada seluruh pengujian disajikan pada Tabel 2.

Tabel 2. Perbandingan Hasil Evaluasi Model Menggunakan 10-Fold Cross Validation

Algoritma Klasifikasi	Akurasi (%)	Presisi (%)	Recall (%)	F1-Score (%)	Waktu Latih (s)
Support Vector Machine (Linear)	99,80%	100,00%	99,60%	99,80%	0,14 s
Multinomial Naïve Bayes	94,00%	93,90%	94,20%	94,00%	0,03 s

Hasil pengujian pada Tabel 2 membuktikan bahwa algoritma Support Vector Machine mencapai performa yang superior dibandingkan Multinomial Naïve Bayes pada seluruh parameter evaluasi. SVM memperoleh tingkat akurasi 99,80% dengan F1-Score 99,80%, mengungguli Naïve Bayes yang mencatat akurasi 94,00% dan F1-Score 94,00%. Meskipun Naïve Bayes unggul dari segi kecepatan waktu pelatihan (0,03 detik berbanding 0,14 detik), selisih waktu tersebut sangat tidak signifikan pada aplikasi praktis penambahan teks skala menengah, sehingga keunggulan akurasi SVM menjadikannya pilihan model yang jauh lebih andal.

3.3 Analisis Matriks Konfusi (Confusion Matrix)

Untuk memahami pola klasifikasi dan kesalahan prediksi (error analysis) secara terperinci, akumulasi nilai matriks konfusi dari keseluruhan 10 fold pengujian (total 1.000 data ulasan) dipetakan pada Tabel 3.

Tabel 3. Matriks Konfusi Akumulatif Pengujian SVM dan Naïve Bayes

Algoritma	Kelas Aktual	Prediksi Negatif	Prediksi Positif	Tingkat Kesalahan
Support Vector Machine	Aktual Negatif (500)	500 (True Negative)	0 (False Positive)	0,0%
	Aktual Positif (500)	2 (False Negative)	498 (True Positive)	0,4%
Multinomial Naïve Bayes	Aktual Negatif (500)	469 (True Negative)	31 (False Positive)	6,2%
	Aktual Positif (500)	29 (False Negative)	471 (True Positive)	5,8%

Berdasarkan rincian Tabel 3, model Support Vector Machine hanya menghasilkan 2 kesalahan klasifikasi (False Negative) dari 1.000 ulasan yang diuji, tanpa mencatatkan satu pun kesalahan False Positive. Kedua ulasan yang salah diklasifikasikan tersebut merupakan ulasan berstruktur kalimat majemuk pertentangan yang memuat kosakata negatif kuat pada klausa anak kalimat (contoh: 'fitur unggah dokumen lancar walau respon admin agak lambat'). Di lain pihak, model Multinomial Naïve Bayes menghasilkan 60 kesalahan klasifikasi (31 False Positive dan 29 False Negative). Tingginya tingkat kesalahan pada Naïve Bayes disebabkan oleh kelemahan inheren asumsi independensi fitur (conditional independence assumption), di mana model memperlakukan setiap kata secara terpisah tanpa memperhitungkan relasi semantik antarkata dalam kalimat bertingkat.

3.4 Pembahasan dan Implikasi Manajerial untuk Pengembangan Layanan Web

Keberhasilan algoritma Support Vector Machine dalam memetakan polaritas opini ulasan publik memiliki implikasi manajerial yang sangat berharga bagi instansi penyelenggara pelayanan publik. Hasil analisis sentimen ini dapat diintegrasikan secara otomatis ke dalam dasbor manajemen sistem informasi eksekutif (Executive Information System) untuk memantau indeks kepuasan masyarakat secara real-time. Berdasarkan temuan bobot fitur negatif pada penelitian ini, rekomendasi aksi yang mendesak dilakukan oleh tim teknis pengembang web mencakup:

1. Optimalisasi Infrastruktur Server: Mengatasi kendala server down pada jam kerja sibuk melalui penerapan load balancing, penskalaan server berbasis komputasi awan (cloud autoscaling), serta kompresi aset web statis.
2. Peningkatan Mekanisme Pengunggahan Lampiran: Menyediakan fitur kompresi gambar dan dokumen otomatis pada sisi peramban klien (client-side auto-compression) sehingga batas ukuran berkas tidak menghambat pemohon.



3. Redesain Antarmuka Berorientasi Pengguna: Menyederhanakan alur formulir berbelit menjadi beberapa tahapan bertahap (multi-step wizard form) yang dilengkapi indikator persentase pengisian data dan penyimpanan otomatis draft formulir.

4. Transparansi Pelacakan Dokumen: Menghubungkan modul notifikasi real-time berbasis API WhatsApp atau SMS Gateway resmi guna menginformasikan status verifikasi berkas secara akurat kepada masyarakat.

IV. KESIMPULAN DAN SARAN

Penelitian ini telah berhasil membandingkan performa algoritma Support Vector Machine dan Multinomial Naïve Bayes dalam mengklasifikasikan sentimen pengguna aplikasi layanan publik berbasis web di Indonesia. Berdasarkan hasil pengujian empiris menggunakan 10-fold cross-validation pada 1.000 ulasan seimbang, algoritma Support Vector Machine dengan kernel linear terbukti unggul secara signifikan dengan tingkat akurasi 99,80%, presisi 100,00%, recall 99,60%, dan F1-score 99,80%, dibandingkan Multinomial Naïve Bayes yang memperoleh akurasi 94,00%, presisi 93,90%, recall 94,20%, dan F1-score 94,00%. Kemampuan SVM dalam memaksimalkan margin bidang pemisah memungkinkan pembedaan polaritas teks yang akurat bahkan pada ulasan dengan kalimat majemuk bertingkat. Temuan riset ini mengidentifikasi bahwa stabilitas sistem saat mengunggah berkas, kecepatan verifikasi, dan transparansi pelacakan permohonan merupakan determinan utama kepuasan masyarakat terhadap sistem informasi layanan publik.

Untuk pengembangan penelitian selanjutnya, disarankan untuk memperluas jangkauan dataset mencakup berbagai platform media sosial seperti X (Twitter) dan ulasan Google Maps, serta mengeksplorasi integrasi model berbasis arsitektur Transformer mutakhir seperti IndoBERT (Devlin et al., 2019) guna menangani ambiguitas semantik dan gaya bahasa daerah yang lebih beragam.

DAFTAR PUSTAKA

- Akriyono, R. V., Rumambi, G. L., Tikoalu, J., Tuilan, V. B. C., & Aktama, G. E. (2025). Penggunaan Bootstrap dalam penerapan responsive web design studi kasus: Template website SMS masking. *Jurnal Logika*, 13(1), 23–36.
- Aktama, G. E., Dali, I. H., Getroyda, M. B., & Baeruma, G. E. (2025). Perancangan sistem informasi website marketing digital pada Universitas Parna Raya. *Jurnal Logika*, 13(1), 37–50.
- Anggraini, F., Defit, S., & Nurcahyo, G. W. (2020). Model prediksi sentimen masyarakat menggunakan pendekatan supervised machine learning. *Jurnal KomTekInfo*, 7(3), 210–218. <https://doi.org/10.35134/komtekinfo.v7i3.88>
- Cahyono, Y., & Suharyanto, S. (2022). Pemanfaatan algoritma machine learning dalam klasifikasi sentimen opini publik pada media sosial: Suatu tinjauan literatur sistematis. *Jurnal Sains dan Informatika*, 8(2), 156–166. <https://doi.org/10.34128/jsi.v8i2.482>
- Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (pp. 4171–4186). Association for Computational Linguistics. <https://doi.org/10.18653/v1/N19-1423>
- Fahrezi, M. R. A., Santoso, H. B., & Isal, R. Y. K. (2022). Evaluasi kepuasan pengguna terhadap sistem informasi layanan terpadu berdasarkan aspek usability dan sentimen pengguna. *Jurnal Sistem Informasi (Journal of Information Systems)*, 18(1), 67–80. <https://doi.org/10.21609/jsi.v18i1.1118>
- Iriananda, S. W., & Rangkuti, A. H. (2021). Analisis sentimen terhadap penyelenggaraan pelayanan publik berbasis web menggunakan lexicon-based dan SVM. *Jurnal Nasional Teknik Elektro dan Teknologi Informasi (JNTETI)*, 10(4), 380–388. <https://doi.org/10.22146/jnteti.v10i4.2384>



- Kurniawan, D., & Utami, T. H. (2022). Perbandingan metode Naive Bayes dan SVM untuk analisis sentimen terhadap layanan e-government. *Jurnal RESTIKOM: Riset Teknik Informatika dan Komputer*, 4(3), 185–194.
- Marlina, L., & Hasan, M. A. (2022). Komparasi algoritma Support Vector Machine dan K-Nearest Neighbor untuk analisis sentimen kepuasan pengguna aplikasi web publik. *Jurnal Teknologi Informasi dan Terapan (J-TIT)*, 9(1), 35–42. <https://doi.org/10.25047/jtit.v9i1.261>
- Merawati, N. L. P., Candiasa, I. M., & Gunadi, I. G. A. (2022). Penerapan k-fold cross validation dalam optimalisasi hyperparameter algoritma SVM untuk text classification. *Jurnal Nasional Pendidikan Teknik Informatika (JANAPATI)*, 11(2), 115–124. <https://doi.org/10.23887/janapati.v11i2.47612>
- Pratama, A., Adiya, M. H., & Rini, D. P. (2023). Analisis sentimen kebijakan publik menggunakan algoritma Support Vector Machine dan Naïve Bayes. *Jurnal Sistem Informasi dan Informatika (SIMIKA)*, 6(1), 45–56. <https://doi.org/10.47065/simika.v6i1.289>
- Putra, I. K. G. D., & Rusjyanthi, N. K. D. (2021). Pengaruh seleksi fitur dan pembobotan TF-IDF pada klasifikasi teks ulasan bahasa Indonesia. *Majalah Ilmiah Teknologi Elektro*, 20(1), 89–98. <https://doi.org/10.24843/MITE.2021.v20i01.P11>
- Rahmad, C., Rahutomo, F., & Susetyo, Y. A. (2022). Komparasi kinerja algoritma klasifikasi teks pada ulasan aplikasi pelayanan publik mobile. *Jurnal Buana Informatika*, 13(2), 121–131. <https://doi.org/10.24002/jbi.v13i2.5641>
- Santoso, B., & Nugroho, R. A. (2021). Penerapan text mining untuk klasifikasi opini masyarakat terhadap pelayanan publik daring. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 5(4), 712–720. <https://doi.org/10.29207/resti.v5i4.3245>
- Setiawan, H., Putra, A. S., & Al-Ghani, M. Y. (2021). Text mining and sentiment analysis of e-government services using Support Vector Machines. *International Journal of Advanced Computer Science and Applications (IJACSA)*, 12(8), 550–558. <https://doi.org/10.14569/IJACSA.2021.0120864>
- Suryani, E., Hendrawan, R., & Hidayatullah, A. F. (2022). Implementasi pembobotan TF-IDF dan algoritma Support Vector Machine dalam analisis sentimen pengguna sistem web. *Jurnal Ilmiah Teknologi Sistem Informasi*, 8(1), 28–37.
- Wahyudi, K. W., Sukarsa, I. M., & Piarsa, I. N. (2021). Sistem penilaian layanan publik berbasis analisis sentimen dengan klasifikasi Naive Bayes. *Jurnal Ilmiah Merpati (Menara Penelitian Akademika Teknologi Informasi)*, 9(2), 142–153. <https://doi.org/10.24843/JIM.2021.v09.i02.p05>
- Wahyuningrum, T. (2021). Mengukur kepuasan pengguna sistem informasi menggunakan metrik usability dan analisis teks ulasan. *Jurnal Edukasi dan Penelitian Informatika (JEPIN)*, 7(3), 332–340. <https://doi.org/10.26418/jp.v7i3.49812>