Journal of Science and Social Research ISSN 2615 – 4307 (Print)
August 2026, IX (4): 6792 – 6800 ISSN 2615 – 3262 (Online)
Available online at http://jurnal.goretanpena.com/index.php/JSSR

ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI IDENTITAS KEPENDUDUKAN DIGITAL MENGGUNAKAN SUPPORT VECTOR MACHINE RANDOM FOREST

Atiqa Auliana Fitri[1], Yuhandri[2], Rini Sovia[3]
Universitas Putra Indonesia YPTK Padang, Sumatera Barat
e-mail: [1]atiqaaulianaf@gmail.com, [2]yuhandri.yunus@gmail.com,
[3]rinisovia2020@gmail.com

Abstract: *The Digital Population Identity (IKD) application allows citizens to access population documents electronically. Service evaluations can be found in user reviews on the Google Play Store, but manual analysis is challenging due to their large and unstructured volume. This study analyzes the sentiment of IKD user reviews and compares the effectiveness of Support Vector Machine (SVM) and Random Forest. Reviews were collected through web scraping, labeled positive or negative by specialists, and preprocessed (cleaning, case folding, tokenizing, normalization, stopword removal, stemming) before being converted into numerical features using TF-IDF weighting. The weighted features were grouped using K-Means clustering to map sentiment tendencies prior to classification. SVM classified sentiment by identifying a maximum-margin hyperplane across classes, while Random Forest combined the predictions of numerous decision trees through majority voting. Both models were trained and tested with a 70:30 stratified split and evaluated using a confusion matrix. From 9,243 scraped reviews collected between June 2025 and June 2026, 5,714 remained after cleaning. SVM achieved 93.64% accuracy with weighted precision, recall, and F1-score of 0.94, while Random Forest achieved 93.41% accuracy with an F1-score of 0.93. Sentiment was dominated by the negative class at 75.5%, mainly complaints about application performance. These results provide a data-driven basis for the Directorate General of Population and Civil Registration to prioritize feature improvements and enhance IKD service quality.*

Keyword: *sentiment analysis; IKD; SVM; Random Forest*

Abstrak: Aplikasi Identitas Kependudukan Digital (IKD) memungkinkan masyarakat mengakses dokumen kependudukan secara elektronik. Ulasan pengguna di Google Play Store memuat penilaian layanan, namun jumlahnya besar dan tidak terstruktur sehingga sulit dianalisis secara manual. Penelitian ini menganalisis sentimen ulasan pengguna IKD dan membandingkan kinerja Support Vector Machine (SVM) dengan Random Forest. Data ulasan dikumpulkan melalui web scraping, diberi label positif dan negatif oleh pakar, lalu dipraproses (cleaning, case folding, tokenizing, normalisasi, stopword removal, stemming) sebelum diubah menjadi fitur numerik dengan pembobotan TF-IDF. Fitur hasil pembobotan dikelompokkan dengan K-Means Clustering untuk memetakan kecenderungan sentimen sebelum klasifikasi. SVM mengklasifikasikan sentimen dengan mencari hyperplane bermargin maksimum antar kelas, sedangkan Random Forest menggabungkan prediksi banyak decision tree melalui majority voting. Kedua model dilatih dan diuji dengan pembagian data 70:30 stratified serta dievaluasi menggunakan confusion matrix. Dari 9.243 ulasan hasil scraping pada rentang Juni 2025 sampai Juni 2026, tersisa 5.714 ulasan setelah pembersihan. SVM memperoleh akurasi 93,64% dengan precision, recall, dan F1-score tertimbang 0,94, sedangkan Random Forest memperoleh akurasi 93,41% dengan F1-score 0,93. Sentimen didominasi kelas negatif sebesar 75,5%, terutama keluhan pada kinerja aplikasi. Hasil penelitian ini menjadi dasar evaluasi berbasis data bagi Direktorat Jenderal Kependudukan dan Pencatatan Sipil untuk memprioritaskan perbaikan fitur dan meningkatkan kualitas layanan IKD.

Kata kunci: analisis sentimen; IKD; SVM; Random Forest

PENDAHULUAN

Administrasi kependudukan merupakan salah satu pilar utama pelayanan publik di Indonesia. Sebagai bagian dari transformasi digital, Direktorat Jenderal Kependudukan dan Pencatatan Sipil meluncurkan aplikasi Identitas Kependudukan Digital (IKD) pada tahun 2022 untuk memungkinkan masyarakat mengakses dokumen kependudukan secara elektronik tanpa harus membawa dokumen fisik (Pratmanto et al., 2023). Sejak diluncurkan, IKD memperoleh beragam tanggapan pengguna di Google Play Store, sebagaimana aplikasi pemerintah lain seperti Sirekap (Tarigan et al., 2025). Ulasan semacam ini umumnya berupa teks tidak terstruktur dalam jumlah besar sehingga sulit dianalisis manual dan menuntut metode otomatis untuk mengklasifikasikan sentimen secara cepat, objektif, dan berskala besar (Garuda et al., 2025).

Ulasan pada platform distribusi aplikasi menjadi sumber evaluasi penting bagi layanan publik (Arifin et al., 2025). Volume ulasan yang tinggi menuntut pendekatan machine learning agar opini pengguna dapat dipetakan secara efisien, seperti pada ulasan aplikasi Halo BCA yang berjumlah ribuan (Adi et al., 2024), karena analisis sentimen terbukti efektif mengubah umpan balik pengguna menjadi informasi pendukung keputusan berbasis data (Cahyani & Prasetyaningrum, 2026).

Analisis sentimen umumnya diawali dengan pengumpulan dan pelabelan ulasan ke dalam kelas sentimen positif dan negatif (Pratmanto et al., 2023), dilanjutkan tahap praproses untuk membersihkan dan menyeragamkan data sebelum ekstraksi fitur dan klasifikasi (Anggraini et al., 2025), lalu direpresentasikan secara numerik menggunakan pembobotan Term Frequency-Inverse Document Frequency (TF-IDF) sebagai masukan bagi algoritma klasifikasi, meskipun teknik ini memiliki keterbatasan dibandingkan word embedding dalam menangkap makna kontekstual pada ulasan berbahasa Indonesia (Idris et al., 2025).

Support Vector Machine (SVM) bekerja dengan membentuk hyperplane bermargin maksimum untuk memisahkan kelas sentimen, sehingga andal mengolah data teks berdimensi tinggi hasil pembobotan TF-IDF dengan risiko overfitting yang relatif rendah (Garuda et al., 2025). Random Forest menerapkan pendekatan ensemble learning yang menggabungkan banyak decision tree melalui voting untuk prediksi yang lebih stabil (Surya & Yamasari, 2026), meskipun SVM terbukti lebih konsisten pada validasi silang (Wiyanto et al., 2025). Studi komparatif pada ulasan Shopee menunjukkan SVM mencapai akurasi 91,77% dan mengungguli Random Forest, dengan perbedaan kinerja yang teruji signifikan secara statistik menggunakan ANOVA (Susanto et al., 2025).

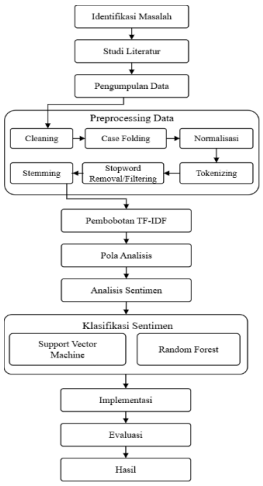
Perbandingan SVM dan Random Forest telah diterapkan pada berbagai domain, seperti quick count Pemilu 2024 dengan performa kompetitif pada data tidak seimbang (Septiana & Alita, 2024), pada isu kesehatan diabetes untuk mencari model klasifikasi terbaik (Hendrastuty, 2025), serta pada isu bahan bakar E10 di TikTok dan kebijakan PSBB untuk menilai algoritma yang paling adaptif terhadap karakteristik data ulasan (Nurfirdaus et al., 2026), (Adrian et al., 2021). Hasil-hasil tersebut menegaskan relevansi kedua algoritma untuk ulasan aplikasi layanan publik dengan karakteristik data serupa, yaitu opini yang bervariasi, tidak baku, dan sering memuat singkatan (Adi et al., 2024).

Meskipun kedua algoritma telah banyak digunakan, studi komparatif yang secara khusus mengukur kinerjanya pada ulasan aplikasi IKD masih terbatas,

karena penelitian terdahulu pada IKD lebih banyak memakai Naïve Bayes dan K-Nearest Neighbor. Berdasarkan kondisi tersebut, penelitian ini bertujuan menganalisis sentimen ulasan pengguna aplikasi IKD sekaligus membandingkan kinerja SVM dengan Random Forest, dengan tahapan pengumpulan data melalui web scraping, pelabelan, praproses, dan pembobotan TF-IDF sebelum klasifikasi. Kinerja kedua algoritma dibandingkan berdasarkan akurasi, precision, recall, dan F1-score untuk menentukan model terbaik sekaligus memetakan aspek layanan yang paling banyak dikeluhkan pengguna, sebagai dasar evaluasi bagi Direktorat Jenderal Kependudukan dan Pencatatan Sipil dalam perbaikan layanan IKD.

METODE

Metode penelitian menggunakan pendekatan data mining dan text mining untuk menganalisis sentimen ulasan pengguna aplikasi Identitas Kependudukan Digital (IKD) yang diperoleh dari Google Play Store. Tahapan penelitian disusun secara terstruktur mulai dari identifikasi masalah, studi literatur, pengumpulan data, praproses, pembobotan TF-IDF, klasifikasi menggunakan SVM dan Random Forest, hingga evaluasi hasil model, sebagaimana ditunjukkan pada Gambar 1.



Gambar 1 Kerangka Penelitian

Pengumpulan Data

Data penelitian berupa ulasan pengguna aplikasi IKD yang dikumpulkan dari Google Play Store melalui teknik web scraping berbasis bahasa pemrograman Python pada lingkungan Google Colaboratory. Setiap ulasan memuat atribut username, waktu unggah, isi ulasan, dan label sentimen hasil pelabelan manual oleh pakar. Pengumpulan data dilakukan pada rentang Juni 2025 sampai Juni 2026 dan diperoleh 9.243 ulasan. Setelah penghapusan 1.842 ulasan duplikat tersisa 7.401 ulasan, lalu penyaringan berdasarkan variabel penelitian dan pembersihan teks menyisakan 5.714 ulasan yang digunakan pada tahap pemodelan.

Text Mining

Pada penelitian ini, text mining diterapkan melalui tahap praproses berupa cleaning, case folding, normalisasi, tokenizing, stopword removal, dan stemming untuk membersihkan noise dan menyeragamkan data sebelum dianalisis (Anggraini et al., 2025).

Term Frequency-Inverse Document Frequency (TF-IDF)

Nilai *Term Frequency* (TF) dihitung dengan membagi jumlah kemunculan sebuah kata dengan total kata dalam dokumen, seperti pada Persamaan 1.

$$tf(t,d) = \frac{f_{t,d}}{\sum_{k \in d} f_{k,d}} \quad (1)$$

dihitung dengan Persamaan 2.

$$idf(t) = log\ \frac{N}{n_t} \quad (2)$$

$$w_{t,d} = tf(t,d) \times idf(t) \quad (3)$$

Support Vector Machine (SVM)

Bidang pemisah (hyperplane) pada SVM dinyatakan dengan Persamaan 4.

$$f(x) = w \cdot x + b \quad (4)$$

$$K(x_i, x_j) = x_i \cdot x_j \quad (5)$$

Proses pelatihan dilakukan menggunakan metode *Sequential Minimal*

Optimization (SMO). Matrix Hessian dihitung dari perkalian nilai kernel dan label kelas melalui Persamaan 6 kemudian nilai error setiap data diperoleh dengan Persamaan 7.

$$D_{ij} = y_i y_j \left(K(x_i, x_j) + \lambda^2 \right) \quad (6)$$

$$E_i = \sum_j \alpha_j D_{ij} \quad (7)$$

$$\delta\alpha_i = min(max[\gamma(1 - E_i), -\alpha_i], C - \alpha_i) \quad (8)$$

$$\delta\alpha_i = min(max[\gamma(1 - E_i), -\alpha_i], C - \alpha_i) \quad (9)$$

$$b = -\frac{1}{2} \left(w \cdot x^+ + w \cdot x^- \right) \quad (10)$$

$$f(x) = \sum_i \alpha_i y_i \, K(x_i, x) + b \quad (11)$$

Random Forest

Pemilihan titik pemisah pada setiap pohon Random Forest menggunakan kriteria Gini pada Persamaan (12), sedangkan ambang batas pemisah antar nilai fitur ditentukan dengan Persamaan 13 (Adrian et al., 2021)

$$Gini(t) = 1 - \sum_{j=1}^{J} {p_j}^2 \quad (12)$$

$$Threshold = \frac{a + b}{2} \quad (13)$$

Setiap pohon dilatih pada subset acak data dan fitur untuk mengurangi varians dan overfitting (Adrian et al., 2021). Hasil prediksi seluruh pohon digabungkan melalui majority voting seperti pada Persamaan 14 (Nurfirdaus et al., 2026)

$$\hat{y} = mode\{ h_1(x), h_2(x), \dots, h_T(x) \} \quad (14)$$

dengan p_j adalah proporsi kelas j pada node, a dan b merupakan dua nilai fitur berurutan, ŷ adalah kelas hasil prediksi akhir, h_t(x) yakni prediksi pohon ke-t, dan T merupakan jumlah pohon.

HASIL DAN PEMBAHASAN

Bagian ini memaparkan hasil penerapan analisis sentimen pada ulasan pengguna aplikasi Identitas Kependudukan Digital (IKD), mulai dari deskripsi dataset, tahap praproses dan pembobotan TF-IDF, hasil klasifikasi Support Vector Machine (SVM) dan Random Forest, hingga evaluasi kinerja kedua model.

Dataset

Dataset yang digunakan pada tahap pemodelan berjumlah 5.714 ulasan hasil praproses, dengan distribusi label didominasi kelas negatif sebanyak 4.312 ulasan (75,5%) dan positif 1.402 ulasan (24,5%). Sampel data ulasan ditampilkan pada Tabel 1.

Tabel 1 Data Ulasan Aplikasi IKD

No	Username	Content	Sentimen
1	ata dani	Saat dibutuhkan buat checkin malah trouble, sering ke logout dan harus selalu login ulang	Negatif
2	Prabu Sedah	makin bagus	Positif
3	Yusuf Zxuan	eror terus gabisa login	Negatif

Tabel 1 memperlihatkan bentuk data ulasan sebelum diproses, yang berupa teks bebas dengan gaya bahasa sehari-hari, singkatan, dan penulisan tidak baku. Label sentimen pada tahap ini menjadi acuan kebenaran untuk melatih dan menguji model. Sebelum masuk ke tahap klasifikasi, data ulasan diproses melalui praproses dan pembobotan fitur agar dapat dikenali oleh algoritma.

Analisis dan Preprocessing

Data ulasan diproses melalui tahap praproses untuk mengurangi ketidakkonsistenan penulisan sebelum ekstraksi fitur. Tahap cleaning menghapus URL, mention, tagar, angka, tanda baca, emoji, dan spasi berlebih, sekaligus

membuang ulasan duplikat. Case folding menyeragamkan seluruh teks menjadi huruf kecil. Tokenizing memecah kalimat menjadi kumpulan kata. Normalisasi mengubah kata tidak baku dan singkatan ke bentuk baku, seperti "sya" menjadi "saya" dan "pkai" menjadi "pakai". Stopword removal membuang kata umum yang tidak berpengaruh terhadap penentuan sentimen. Stemming mengembalikan kata berimbuhan ke bentuk dasarnya, seperti "aplikasinya" menjadi "aplikasi". Setelah seluruh tahap praproses dan penyaringan berdasarkan variabel penelitian, diperoleh 5.714 ulasan yang digunakan pada tahap analisis berikutnya.

Tabel 2 Hasil Praproses Data Ulasan

content	cleaning	case_ folding	tokenizing	normalisasi	stopword	stemming
Saat dibutuhkan buat checkin …	Saat dibutuh kan buat checkin …	saat dibutuh kan buat checkin …	[saat, dibutuh kan, buat, checkin, …]	[saat, dibutuhkan, checkin, …]	[dibutuhkan, checkin, login, ulang, …]	[butuh, checkin, login, ulang, …]
eror terus gabisa login	eror terus gabisa login	Eror terus gabisa login	[eror, terus, gabisa, login]	[error, terus, tidak bisa, login]	[error, tidak bisa, login]	[error, tidak, bisa, login]
Proses cetak ulang KTP …	Proses cetak ulang KTP …	proses cetak ulang ktp …	[proses, cetak, ulang, ktp, …]	[proses, cetak, ulang, ktp, …]	[proses, cetak, ulang, ktp, …]	[proses, cetak, ulang, ktp, …]

Teks hasil praproses selanjutnya diubah menjadi representasi numerik menggunakan pembobotan TF-IDF. Nilai TF dihitung dari frekuensi kemunculan kata pada dokumen, nilai IDF dihitung dari sebaran kata pada seluruh dokumen, dan bobot akhir diperoleh dari perkalian keduanya. Pembobotan dilakukan berdasarkan dua kategori variabel penelitian, yaitu performa aplikasi (X1) dan kemudahan pengguna (X2). Hasil pembobotan dipetakan ke variabel yang sesuai untuk melihat pola kemunculan kata kunci pada tiap aspek. Pola analisis data ditampilkan pada Tabel 3.

Tabel 3 Pola Analisis Data Sampel Berdasarkan Variabel

Data	Performa Aplikasi (X1)	Kemudahan Pengguna (X2)
D1	1,2930	0,9779
D2	1,1609	0,0000
D3	0,0000	0,8265
D4	0,3271	1,2456
D5	1,6060	1,2084
…	…	…
D5717	1,5464	1,2456
D5718	0,0000	0,0000
D5719	1,0637	0,0000
D5720	0,3271	1,2818
D5721	0,0000	1,1583

K-Means Clustering

Bobot TF-IDF kedua variabel selanjutnya dikelompokkan menggunakan algoritma K-Means Clustering ke dalam dua cluster untuk melihat pola kecenderungan sentimen sebelum klasifikasi. Hasil pengelompokan ditampilkan pada Tabel 4.

Tabel 4 K-Means Clustering

Dokumen	Kinerja Aplikasi	Kemudahan Pengguna	Cluster	Sentimen
D1	1,2930	0,9779	1	Negatif
D2	1,1609	0,0000	0	Negatif

Dokumen	Kinerja Aplikasi	Kemudahan Pengguna	Cluster	Sentimen
D3	0,0000	0,8265	1	Positif
D4	0,3271	1,2456	1	Negatif
D5	1,6060	1,2084	1	Negatif
…	…	…	…	…
D5717	1,5464	1,2456	1	Negatif
D5718	0,0000	0,0000	0	Negatif
D5719	1,0637	0,0000	0	Negatif
D5720	0,3271	1,2818	1	Negatif
D5721	0,0000	1,1583	1	Negatif

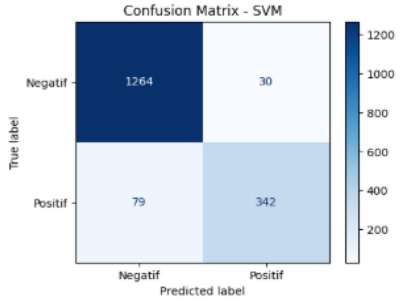
Algoritma K-Means mengelompokkan data ke dalam dua cluster, yaitu Cluster 0 sebanyak 2.912 dokumen dengan kecenderungan sentimen positif dan Cluster 1 sebanyak 2.809 dokumen dengan kecenderungan sentimen negatif. Titik centroid kedua cluster pada variabel kinerja aplikasi dan kemudahan pengguna berturut-turut adalah (0,011; -0,928) dan (-0,011; 0,962), dengan proses iterasi mencapai konvergen pada iterasi ke-7. Hasil pengelompokan ini memperkuat pelabelan pakar dan menjadi dasar bagi klasifikasi sentimen menggunakan Support Vector Machine dan Random Forest.

Support Vector Machine (SVM)

Data hasil pembobotan TF-IDF sebanyak 5.714 ulasan dibagi menjadi data latih dan data uji dengan perbandingan 70:30 secara stratified, sehingga diperoleh 3.999 data latih dan 1.715 data uji. Model Support Vector Machine dilatih pada data latih menggunakan kernel linear, kemudian diuji pada data uji dengan membandingkan label prediksi terhadap label sebenarnya. Dari 1.715 data uji, model mengklasifikasikan 1.606 ulasan dengan benar dan 109 ulasan keliru, sehingga diperoleh akurasi sebesar 93,64%. Hasil klasifikasi model SVM disajikan pada Tabel 5, confusion matrix pada Gambar 2, dan laporan klasifikasinya pada Tabel 6.

Tabel 5 Hasil Klasifikasi Model SVM

No	Label Aktual	Label Prediksi	Keterangan
1	Positif	Positif	Benar
2	Negatif	Positif	Salah
3	Negatif	Negatif	Benar
4	Negatif	Negatif	Benar
5	Negatif	Negatif	Benar



Gambar 2 Confusion Matrix Model SVM

Tabel 6 Laporan Klasifikasi Model SVM

Kelas	Precision	Recall	F1-Score	Support
Negatif	0,94	0,98	0,96	1294
Positif	0,92	0,81	0,86	421
Akurasi			0,94	1715
Rata-rata Makro	0,93	0,89	0,91	1715
Rata-rata Tertimbang	0,94	0,94	0,94	1715

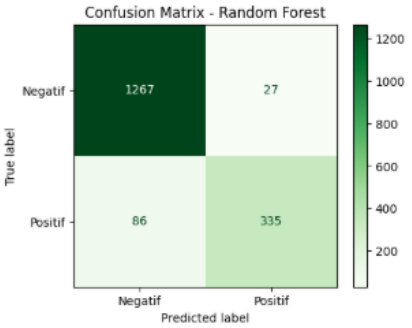
Random Forest

Model Random Forest diuji pada 1.715 data uji yang sama dengan pembagian data 70:30. Model menggabungkan prediksi sejumlah decision tree melalui mekanisme majority

voting untuk menentukan kelas akhir setiap ulasan. Dari 1.715 data uji, model mengklasifikasikan 1.602 ulasan dengan benar dan 113 ulasan keliru, sehingga diperoleh akurasi sebesar 93,41%. Hasil klasifikasi model Random Forest disajikan pada Tabel 7, confusion matrix pada Gambar 3, dan laporan klasifikasinya pada Tabel 8.

Tabel 7 Hasil Klasifikasi Model Random Forest

No	Label Aktual	Label Prediksi	Keterangan
1	Positif	Positif	Benar
2	Negatif	Negatif	Benar
3	Negatif	Negatif	Benar
4	Negatif	Negatif	Benar
5	Negatif	Negatif	Benar



Gambar 3 Confusion Matrix Model Random Forest

Gambar 3 menampilkan confusion matrix model Random Forest pada data uji. Polanya serupa dengan SVM, dengan kesalahan pada kelas positif sedikit lebih banyak sehingga akurasinya sedikit lebih rendah.

Tabel 8 Laporan Klasifikasi Model Random Forest

Kelas	Precision	Recall	F1-Score	Support
Negatif	0,94	0,98	0,96	1294
Positif	0,93	0,80	0,86	421
Akurasi			0,93	1715
Rata-rata Makro	0,93	0,89	0,91	1715
Rata-rata Tertimbang	0,93	0,93	0,93	1715

Berdasarkan Tabel 8, model Random Forest menghasilkan kinerja yang mendekati SVM, dengan recall pada kelas positif sebesar 0,80 yang kembali dipengaruhi distribusi data tidak seimbang, dan rata-rata tertimbang precision, recall, dan F1-score sebesar 0,93.

Evaluasi

Kinerja kedua model dibandingkan untuk menentukan algoritma dengan performa terbaik pada klasifikasi sentimen ulasan aplikasi IKD. Perbandingan akurasi dan metrik tertimbang kedua model disajikan pada Tabel 9.

Tabel 9 Perbandingan Kinerja Model SVM dan Random Forest

Metrik	SVM	Random Fores
Akurasi	93,64%	93,41%
Precision	0,94	0,93
Recall	0,94	0,93
F1-Score	0,94	0,93

Berdasarkan Tabel 9, model SVM memperoleh akurasi 93,64% sedangkan Random Forest 93,41%, dengan selisih 0,23%. Tingkat error model dihitung dari rasio jumlah data salah terhadap jumlah data uji, sehingga diperoleh error SVM sebesar 6,36% dan Random Forest sebesar 6,59%.

Dari 5.714 ulasan yang dianalisis, sentimen didominasi kelas negatif sebesar 75,5%, terutama keluhan pada aspek performa seperti gangguan sistem dan kegagalan login, sehingga dapat menjadi dasar bagi Direktorat Jenderal Kependudukan dan Pencatatan Sipil dalam memprioritaskan perbaikan. Penelitian ini turut menjawab keterbatasan studi komparatif SVM dan Random Forest pada ulasan aplikasi IKD, sekaligus menunjukkan bahwa kombinasi TF-IDF dengan SVM memberikan kinerja terbaik untuk memetakan persepsi

pengguna.

SIMPULAN

Penelitian ini menganalisis sentimen ulasan pengguna aplikasi IKD di Google Play Store sekaligus membandingkan kinerja SVM dan Random Forest, menggunakan 5.714 ulasan hasil pembersihan dari 9.243 ulasan awal yang diproses melalui praproses, pembobotan TF-IDF, pengelompokan K-Means, dan klasifikasi pada pembagian data 3.999 data latih dan 1.715 data uji. SVM memperoleh akurasi 93,64% dengan F1-score tertimbang 0,94, sedangkan Random Forest memperoleh akurasi 93,41% dengan F1-score 0,93, sehingga SVM menjadi model terbaik dengan selisih akurasi 0,23% dan error lebih rendah. Sentimen ulasan didominasi kelas negatif sebesar 75,5%, terutama keluhan pada gangguan sistem dan kegagalan login, sehingga dapat menjadi dasar evaluasi bagi Direktorat Jenderal Kependudukan dan Pencatatan Sipil dalam memprioritaskan perbaikan layanan IKD.

DAFTAR PUSTAKA

ADDIN Mendeley Bibliography CSL_BIBLIOGRAPHY Adi, S., Mola, S., Lyan, D., Baun, B., Nunes, I. O., Maria, M., & Rani, A. (2024). *ANALISIS SENTIMEN APLIKASI HALO BCA DI GOOGLE PLAY STORE MENGGUNAKAN METODE NAIVE BAYES , SUPPORT VECTOR MACHINE DAN RANDOM FOREST PENDAHULUAN*. *15*(c), 69–79.

Adrian, M. R., Putra, M. P., Rafialdy, M. H., Rakhmawati, N. A., & Informasi, D. S. (2021). *Perbandingan Metode Klasifikasi Random Forest dan SVM Pada Analisis Sentimen PSBB*. *7*(1), 36–40.

Anggraini, P., Informasi, S., Nasional, U., Manila, J. S., Minggu, P., & Selatan, J. (2025). *KOMPARASI NAÏVE BAYES , SUPPORT VECTOR MACHINE , DAN RANDOM FOREST DALAM ANALISIS SENTIMEN*. *9*(3), 4451–4457.

Arifin, M. N., Hamzah, A., Huda, M. A., & Hasanah, N. (2025). *Analysis of Google Play Store User Sentiment Towards Application X Using the SVM Algorithm*. *5*(1), 249–258.

Cahyani, R. D., & Prasetyaningrum, P. T. (2026). *Sentiment Analysis of User Reviews for AI Applications : Evaluating SVM , Logistic Regression , and Random Forest*. *8*(1), 1–27. https://doi.org/10.63158/journalisi.v8i1.1366

Garuda, I., Alva, P., Zuliarso, E., & Stikubank, U. (2025). *SENTIMENT ANALYSIS BASED ON GOOGLE PLAY STORE REVIEWS OF THE*. *8*.

Hendrastuty, N. (2025). *Analisis Sentiment Terhadap Diabetes Menggunakan Algoritma Naïve Bayes , Random Forest , SVM Pada Media Sosial X*. *6*(4), 2469–2479. https://doi.org/10.47065/bits.v6i4.6941

Idris, M., Rifai, A., & Tania, K. D. (2025). *Sentiment Analysis of Tokopedia App Reviews using Machine Learning and Word Embeddings*. *9*(1), 210–219.

Nasional, J., Informasi, S., Cahaya, N., Tahyudin, I., & Shouni, A. (2024). *Perbandingan Algoritma Support Vector Machine , Decision Tree , dan Logistic Regresion Pada Analisis Sentimen Ulasan Aplikasi Netflix*. *02*, 110–117.

Nevrada, N. A., & Syaputra, M. A. (2025). *Sentiment Analysis of Telegram App Reviews on Google Play Store Using the Support Vector Machine (SVM) Algorithm*. *9*(1), 96–105.

Nurfirdaus, R. E., Barata, M. A., & Prastya, I. W. (2026). *Comparison of SVM and Random Forest for TikTok E10 Fuel Sentiment Analysis*. *10*(2), 1426–1437.

Pratmanto, D., Fandi, F., Imaniawan, D., Maarif, V., Studi, P., Komputer, T., Informasi, F. T., Sarana, U. B., Studi, P., Informasi, S., Informasi, F. T., Sarana, U. B., & Pusat, J. (2023). *KEPENDUDUKAN DIGITAL DENGAN METODE*. *2*, 151–161.

Risanindya, T., Purbasari, W., & Riyandari, L. (2025). *Analisis Sentimen Ulasan Pengguna Aplikasi Maxim pada Google Play Store dengan Metode Support Vector Machine (SVM) dan Naïve Bayes Sentiment Analysis of Maxim App User Reviews on Google Play Store with Support Vector Machine (SVM) and Naïve Bayes Methods*. *14*(105), 859–867.

Septiana, I., & Alita, D. (2024). *Perbandingan Random Forest dan SVM dalam Analisis Sentimen Quick Count Pemilu 2024*. *9*(3), 224–233. https://doi.org/10.30591/jpit.v9i3.664

Surya, R. J., & Yamasari, Y. (2026). *Eksplorasi Random Forest Berbasis Ekstraksi Fitur dan Metode Sampling untuk Analisis Data Multi-label Sentimen dan Emosi*. *7*(4), 822–831.

Susanto, E., Purnama, B., Informasi, S., Komputer, I., & Jambi, U. D. (2025). *Komparasi Algoritma SVM dan Random Forest Dalam Sentimen Analisis Review Shopee di Google Play Store Dengan Anova*. *November*.

Tarigan, D. A., Situmorang, Z., & Rosnelly, R. (2025). *ANALISIS SENTIMEN APLIKASI PLAYSTORE SIREKAP 2024 PASCA PILPRES DENGAN PERBANDINGAN METODE SUPPORT VECTOR MACHINE (SVM), SENTIMENT ANALYSIS OF THE SIREKAP 2024 PLAYSTORE APPLICATION POST-PRESIDENTIAL ELECTION WITH COMPARISON OF SUPPORT VECTOR MACHINE (*. *11*(3), 661–670.

Wiyanto, A., Puspasari, S., Astuti, L. W., Komputer, I., Informatika, T., Indo, U., & Mandiri, G. (2025). *Perbandingan Kinerja Algoritma Support Vector Machine dan Random Forest dalam Analisis Sentimen Ulasan Hotel di Kota Palembang pada Google*.