JURNAL KOMPUTER TEKNOLOGI INFORMASI SISTEM KOMPUTER

p-ISSN: 2962-3022 | e-ISSN: 2963-7104
Volume 5 No 2 September 2026
Hal: 1242-1251
http://ejurnal.lkpkaryaprima.id/index.php/juktisi
DOI: https://doi.org/10.62712/juktisi.v5i2.1324

Perbandingan Metode Support Vektor Machine dan Random Forest untuk Analisis Sentimen Ulasan Adapundi dan Kreditpintar di Google Playstore

Lupias[1], Panny Agustia Rahayuningsih[2], Muhammad Rezki[3*]

[1,2,3]Program Studi Informatika, Universitas Bina Sarana Informatika, Pontianak, Kalimantan Barat, Indonesia
Email: 15220518@bsi.ac.id[1], Panny.par@bsi.ac.id[2], muhammad.mdk@bsi.ac.id[3*]
(*Email Corresponding Author: muhammad.mdk@bsi.ac.id[3])
Received: 25 Juni 2026 | Revision: 30 Juni 2026 | Accepted: 1 Juli 2026

Abstrak

Layanan pinjaman online (pinjol) seperti AdaPundi dan KreditPintar menghasilkan ribuan ulasan pengguna di Google Play Store yang bersifat tidak terstruktur, sehingga sulit dianalisis secara manual untuk memahami persepsi publik. Penelitian ini bertujuan membandingkan performa algoritma Support Vector Machine (SVM) dan Random Forest dalam mengklasifikasikan sentimen ulasan kedua aplikasi tersebut, serta mengidentifikasi kata kunci yang membentuk sentimen positif dan negatif. Data dikumpulkan melalui scraping sebanyak 4.998 ulasan, dibersihkan menjadi 2.739 baris, dan diberi label sentimen berdasarkan rating (1–3 negatif, 4–5 positif), menghasilkan proporsi 69,19% positif dan 30,81% negatif. Pra-pemrosesan teks dilakukan melalui sepuluh tahap dengan modifikasi khusus berupa proteksi kata negasi, menghasilkan 2.695 data bersih yang selanjutnya diekstraksi menjadi fitur TF-IDF (matriks 2.695×1.204, sparsity 99,48%). Ketidakseimbangan kelas ditangani melalui class_weight='balanced' tanpa SMOTE. Hasil pengujian pada data uji menunjukkan SVM kernel RBF memperoleh akurasi 87,76%, F1-Score 0,8794, dan MCC 0,7276, mengungguli Random Forest yang memperoleh akurasi 86,64%, F1-Score 0,8670, dan MCC 0,6919, meskipun Random Forest mencatat ROC-AUC lebih tinggi (0,9216 berbanding 0,9077). Validasi 5-fold cross validation menguatkan temuan ini secara konsisten. Perbandingan kedua aplikasi menunjukkan AdaPundi memperoleh sentimen lebih positif (74,4%) dibandingkan KreditPintar (63,8%), dengan kata "tolak" menjadi penanda khas keluhan pada KreditPintar terkait penolakan pengajuan pinjaman.

Kata Kunci: Analisis Sentimen, Support Vector Machine, Random Forest, Pinjaman Online, TF-IDF

Abstract

Online lending (pinjol) services such as AdaPundi and KreditPintar generate thousands of unstructured user reviews on the Google Play Store, making manual analysis impractical for understanding public perception. This study aims to compare the performance of Support Vector Machine (SVM) and Random Forest algorithms in classifying the sentiment of reviews from both applications, as well as to identify keywords that shape positive and negative sentiment. Data were collected through scraping, yielding 4,998 reviews, which were cleaned to 2,739 rows and labeled based on star ratings (1–3 as negative, 4–5 as positive), resulting in a proportion of 69.19% positive and 30.81% negative. Text preprocessing was carried out through ten stages with a special modification to protect negation words, producing 2,695 clean records that were subsequently extracted into TF-IDF features (a 2,695×1,204 matrix with 99.48% sparsity). Class imbalance was handled using class_weight='balanced' instead of SMOTE. Test results show that the RBF-kernel SVM achieved an accuracy of 87.76%, an F1-Score of 0.8794, and an MCC of 0.7276, outperforming Random Forest, which achieved an accuracy of 86.64%, an F1-Score of 0.8670, and an MCC of 0.6919, although Random Forest recorded a higher ROC-AUC (0.9216 versus 0.9077). Five-fold cross-validation results consistently reinforced these findings. The comparison between the two applications shows that AdaPundi obtained more positive sentiment (74.4%) than KreditPintar (63.8%), with the word "tolak" (rejected) emerging as a distinctive marker of complaints on KreditPintar related to loan application rejection.

***Keywords**: Sentiment Analysis, Support Vector Machine, Random Forest, Online Lending, TF-ID*

1. PENDAHULUAN

Dalam beberapa tahun terakhir, pinjaman daring (pinjol) telah menjadi salah satu produk fintech yang paling berpengaruh di masyarakat Indonesia. Inti dari konsepnya adalah sangat sederhana: tidak perlu mendatangi kantor fisik, tanpa perlu menyiapkan banyak dokumen, dan dana dapat diterima hanya dengan beberapa kali klik di ponsel. Kemudahan inilah yang menjadikan pinjol sangat populer, terutama di kalangan generasi muda yang kerap mencari uang secara cepat tanpa mengalami proses administrasi yang rumit. Namun, kenyamanan ini juga memunculkan sisi negatif mulai dari bunga yang dianggap menekan, metode penagihan yang kadang terasa menakutkan, hingga tahapan verifikasi data yang bisa menyulitkan pengguna. AdaPundi dan KreditPintar merupakan dua nama terkenal di antara platform pinjol yang telah mendapatkan izin resmi dari Otoritas Jasa Keuangan (OJK), dan kedua perusahaan ini meninggalkan jejak

JURNAL KOMPUTER TEKNOLOGI INFORMASI SISTEM KOMPUTER
p-ISSN: 2962-3022 | e-ISSN: 2963-7104
Volume 5 No 2 September 2026
Hal: 1242-1251
http://ejurnal.lkpkaryaprima.id/index.php/juktisi
DOI: https://doi.org/10.62712/juktisi.v5i2.1324

ulasan yang sangat banyak di Google Play Store jejak digital yang pada dasarnya mencerminkan bagaimana masyarakat menilai pelayanan mereka.

Sayangnya, gambar tersebut belum terorganisir dengan baik. Ribuan review yang ditulis oleh pengguna tersimpan dalam format teks yang tidak terstruktur, membuatnya terasa mustahil untuk dibaca satu per satu secara manual demi memahami pandangan publik. Masalah ini menjadi lebih rumit akibat adanya bahasa sehari-hari yang khas dalam ulasan penggunaan singkatan, istilah gaul, dan hal yang paling sering terabaikan: struktur kalimat yang mengandung negasi seperti "tidak bagus". Apabila tahap persiapan teks tidak dikelola dengan hati-hati, kata "tidak" dapat saja dihapus sebagai kata yang tidak penting, sehingga kalimat yang awalnya memiliki makna negatif malah bisa dianggap positif oleh sistem. Untuk menyelesaikan permasalahan ini, penelitian ini menggunakan metode text mining berupa analisis sentimen otomatis yang didasarkan pada machine learning, disertai dengan tahap persiapan khusus untuk menjaga kata-kata negasi agar makna asli dari ulasan tetap terpelihara saat kata yang tidak penting dibersihkan.

Sebenarnya, kajian mengenai sentimen dalam ulasan aplikasi keuangan bukanlah tema yang baru banyak pihak telah mengeksplorasinya dalam lima tahun terakhir, namun dengan perspektif yang beragam. Iqbal dan timnya, contohnya, menggunakan algoritma SVM untuk menganalisis sentimen pada lima aplikasi pinjaman online secara bersamaan, termasuk Kredivo dan Easycash, yang menariknya di antara dua aplikasi tersebut adalah Kredit Pintar dan Ada Pundi sendiri [1]. Rata-rata tingkat akurasi yang didapatkan memang "hanya" mencapai 72%, tetapi khusus untuk aplikasi Kredit Pintar, kinerjanya malah melonjak hingga 83% [1]. Penemuan ini cukup menunjukkan bahwa SVM memiliki potensi yang dapat diandalkan dalam konteks pinjaman online namun penelitian tersebut hanya terbatas pada satu algoritma saja, dan belum pernah dibandingkan secara langsung dengan Random Forest dalam suatu kerangka percobaan yang sebanding.

Pertanyaan mengenai "siapa yang lebih unggul" jadi semakin menarik ketika dianalisis lebih dalam, karena jawabannya bisa bervariasi tergantung pada aplikasinya. Dalam kajian aplikasi Halo BCA, Mola dan tim menguji tiga algoritma Naive Bayes, SVM, dan Random Forest dan Random Forest muncul sebagai pemenang dengan tingkat akurasi 90,01%, mengungguli SVM yang mencapai 86,86% [2]. Namun, dalam kajian aplikasi Gojek, Aditya dan rekan-rekan menemukan hasil yang sebaliknya: SVM berhasil mencatat akurasi 96%, sementara Random Forest tertinggal di angka 93% [3]. Pola yang sama juga terlihat di domain lainnya: dalam ulasan aplikasi marketplace Shopee, Suswadi dan Erkamim mencatat bahwa SVM lebih efektif (84,71%) dibandingkan dengan Random Forest (82,21%) [4]; sementara itu, di ulasan aplikasi crypto exchange Indodax, Maulana dan tim justru memilih Random Forest sebagai model utama yang kemudian dioptimalkan lewat tuning hyperparameter [5]. Di sektor layanan publik, Maheri dan tim menunjukkan bahwa SVM lebih akurat (80,76%) dibandingkan Naive Bayes (78,12%) dalam kajian aplikasi M-Paspor [6]. Di sisi lain, pada aplikasi Polri Super App, Susanto melaporkan bahwa SVM dengan kernel linear mencapai akurasi tertinggi 89,67%, mengungguli K-Nearest Neighbor [7]. Perbedaan dan keragaman hasil di berbagai domain ini menunjukkan satu fakta penting: tidak ada aturan baku mengenai dominasi SVM atau Random Forest; sebenarnya, hal ini sangat tergantung pada sifat masing-masing domain serta pola data yang ada. Maka dari itu, klaim mengenai "siapa yang lebih baik" dalam aplikasi pinjaman online tidak dapat langsung diadopsi dari temuan di domain lain perlu dilakukan uji sendiri.

Selain mengenai pemilihan algoritma, pengelolaan kata negasi pada tahap pra-pemrosesan juga mendapat perhatian khusus dalam studi-studi terbaru. Tarecha dan rekan-rekannya mengusulkan metode inversi dan pengurangan skor sentimen berdasarkan cakupan negasi untuk mempertahankan makna negasi dalam kalimat ulasan yang ditulis dalam bahasa Indonesia [8]. Penelitian ini kemudian diperluas oleh Darmayasa dan tim, yang melakukan perbandingan antara dua metode penanganan negasi Next Word Negation dan substitusi antonim pada ulasan dari marketplace Tokopedia dan Lazada, menemukan bahwa kedua metode ini secara konsisten meningkatkan akurasi, baik untuk Naive Bayes (dari 82,94% menjadi 87,64%) maupun untuk SVM (dari 84,70% menjadi 89,41%) [9]. Temuan ini menjadi bukti yang kuat bahwa pengelolaan negasi bukan hanya sebuah rincian teknis, melainkan sebuah komponen yang bisa mempengaruhi akurasi model secara signifikan sehingga sangat relevan untuk diterapkan secara rutin dalam penelitian ini. Satu hal metodologis lain yang perlu diperhatikan adalah masalah ketidakseimbangan kelas yang hampir selalu muncul dalam data ulasan aplikasi. Fitroh dan rekan-rekannya menguji empat skema penanganan baseline, SMOTE, pengalihan kelas, dan kombinasi keduanya pada ulasan ChatGPT, dan menemukan bahwa pengalihan kelas justru lebih konsisten meningkatkan F1-score makro untuk SVM, sementara SMOTE tidak menunjukkan perbaikan yang stabil [10]. Temuan inilah yang menjadi landasan bagi keputusan penelitian ini untuk memilih pengalihan kelas daripada SMOTE dalam menangani ketidakseimbangan kelas pada data AdaPundi dan KreditPintar.

Dari sepuluh kajian sebelumnya tersebut, sejumlah celah dapat dikenali sebagai GAP Analysis untuk penelitian ini. Pertama, satu-satunya penelitian yang berkaitan langsung dengan topik pinjaman online masih terbatas pada satu algoritma saja (SVM), tanpa melakukan pengujian bersamaan dengan Random Forest [1]. Kedua, hasil kompetisi SVM melawan Random Forest dalam berbagai konteks aplikasi lainnya justru menunjukkan hasil yang tidak konsisten dan sangat tergantung pada karakteristik masing-masing domain, sehingga tidak dapat secara langsung diterapkan dalam konteks pinjaman online [2], [3], [4], [5]. Ketiga, hingga saat ini, belum ada studi yang dengan jelas membandingkan dua aplikasi pinjaman online AdaPundi dan KreditPintar dalam satu eksperimen yang setara. Keempat, penelitian mengenai penanganan negasi masih berfokus pada domain marketplace, dan belum ada penerapan dan pengujian manfaatnya secara spesifik dalam konteks ulasan pinjaman online yang banyak diwarnai dengan keluhan bernuansa negasi, seperti "tidak transparan" atau "belum cair" [8], [9]. Kelima, banyak penelitian sebelumnya masih mengandalkan SMOTE untuk menangani ketidakseimbangan kelas, padahal penelitian terbaru menunjukkan bahwa penggunaan class weighting

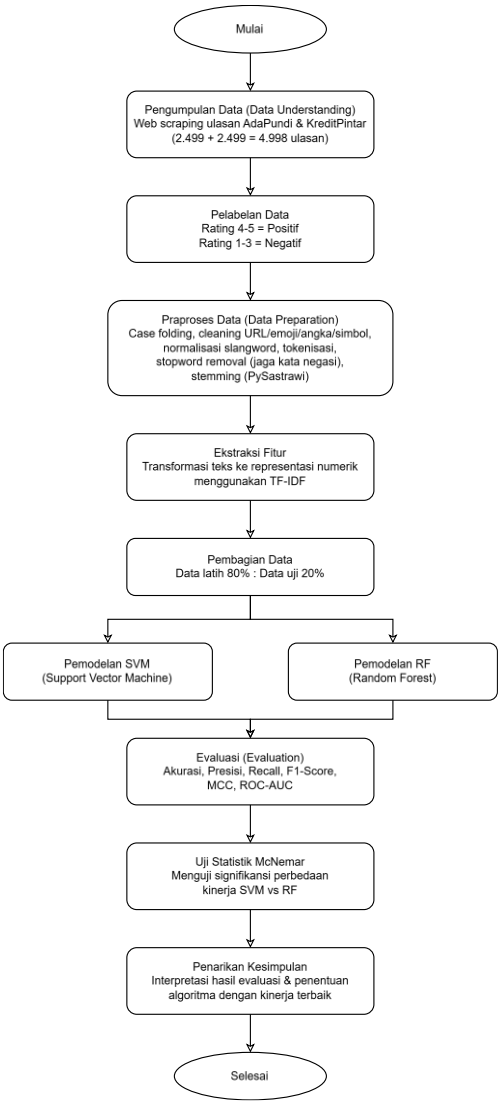
JURNAL KOMPUTER TEKNOLOGI INFORMASI SISTEM KOMPUTER
p-ISSN: 2962-3022 | e-ISSN: 2963-7104
Volume 5 No 2 September 2026
Hal: 1242-1251
http://ejurnal.lkpkaryaprima.id/index.php/juktisi
DOI: https://doi.org/10.62712/juktisi.v5i2.1324

menghasilkan hasil yang lebih konsisten, terutama ketika berhadapan dengan representasi fitur TF-IDF yang bersifat jarang dan berdimensi tinggi [10].

Berdasarkan kesenjangan-kesenjangan yang ada, studi ini ditujukan untuk mencapai tiga sasaran: [1] mengidentifikasi dan membandingkan pendapat pengguna mengenai AdaPundi dan KreditPintar berdasarkan review yang terdapat di Google Play Store; [2] menguji kinerja Support Vector Machine dan Random Forest dalam mengklasifikasikan pandangan terhadap kedua aplikasi tersebut, dengan memanfaatkan fitur TF-IDF sambil menangani masalah ketidakseimbangan kelas melalui pengaturan bobot kelas serta proses pra-pemrosesan yang mempertahankan konteks negasi; dan [3] menyelidiki kata kunci yang paling signifikan dalam membentuk arah pandangan baik yang positif maupun negatif pada masing-masing aplikasi. Diharapkan hasil dari studi ini dapat menjadi dasar pertimbangan bagi penyedia layanan pinjaman online dalam meningkatkan mutu layanannya, sekaligus menjadi referensi metodologis bagi studi serupa di bidang analisis sentimen aplikasi fintech lainnya.

2. METODOLOGI PENELITIAN

Penelitian ini mendefinisikan proses yang terbagi menjadi dua tahap utama, yakni tahap persiapan data dan tahap pemodelan, sehingga setiap langkah bisa dievaluasi kontribusinya terhadap kinerja akhir model dengan cara yang teratur. Seluruh proses penelitian terdiri dari delapan langkah utama seperti yang ditunjukkan pada Gambar 1.



Gambar 1. Alur penelitian

2.1 Pengumpulan Data

Data diperoleh melalui metode scraping di Google Play Store dengan bantuan pustaka google-play-scraper[11], masing-masing dengan 2.499 ulasan untuk aplikasi AdaPundi dan KreditPintar, sehingga total keseluruhan data mentah

JURNAL KOMPUTER TEKNOLOGI INFORMASI SISTEM KOMPUTER
p-ISSN: 2962-3022 | e-ISSN: 2963-7104
Volume 5 No 2 September 2026
Hal: 1242-1251
http://ejurnal.lkpkaryaprima.id/index.php/juktisi
DOI: https://doi.org/10.62712/juktisi.v5i2.1324

mencapai 4.998 baris. Pemilihan Google Play Store sebagai sumber data dipertimbangkan karena platform ini menawarkan ulasan yang sudah diverifikasi dari pengguna yang terdaftar, menjadikannya lebih dapat dipercaya dibandingkan pendapat di media sosial yang sering melibatkan akun anonim atau bot. Atribut yang diekstrak mencakup konten (teks ulasan) dan skor (rating bintang 1–5). Kedua dataset tersebut kemudian disatukan menjadi satu dataframe dengan penambahan kolom sumber untuk menunjukkan asal aplikasi, dilanjutkan dengan proses pembersihan data, yang mencakup penghapusan nilai yang hilang (missing value) pada kolom skor (ditemukan sebanyak 1.518 nilai kosong) serta penghapusan data duplikat pada kolom konten (terdapat 827 baris yang duplikat). Setelah tahap pembersihan, diperoleh 2.739 baris data yang valid untuk digunakan dalam langkah berikutnya.

2.2 Pelabelan Sentimen

Pelabelan sentimen dilakukan secara otomatis dengan berdasarkan pada nilai rating [1], menggunakan kriteria: rating 1, 2, dan 3 dianggap sebagai sentimen negatif, sementara rating 4 dan 5 dinilai sebagai sentimen positif. Metode ini banyak diterapkan dalam studi analisis sentimen yang berlandaskan rating. Hasil dari proses pelabelan menunjukkan bahwa data memiliki komposisi 69,19% positif dan 30,81% negatif (rasio 2,25:1), yang dikategorikan sebagai ketidakseimbangan moderat karena masih berada di bawah ambang batas ekstrem >10:1 [10]. Untuk menyelesaikan masalah ini, penelitian ini menggunakan parameter class_weight='balanced' pada kedua model, didukung oleh teknik stratified split untuk menjaga proporsi kelas di data latih dan data uji, sebagai alternatif dari metode oversampling seperti SMOTE yang dianggap kurang efektif pada representasi fitur TF-IDF yang jarang dan berdimensi tinggi.

2.3 Pra-pemrosesan Teks

Pra-pemrosesan teks dilakukan melalui sepuluh tahap berurutan, yaitu: case folding, penghapusan URL, penghapusan emoji, penghapusan angka dan simbol, normalisasi kata tidak baku (slang) menggunakan kamus sebanyak 153 entri, penghapusan whitespace berlebih, tokenizing, penghapusan stopword (807 kata dari NLTK Bahasa Indonesia dan kamus tambahan), stemming menggunakan pustaka Sastrawi [12], dan penggabungan token kembali menjadi teks bersih. Pada tahap penghapusan stopword, dilakukan modifikasi khusus berupa pengecualian sepuluh kata negasi penting (tidak, bukan, belum, jangan, tanpa, tak, non, anti, kurang, gagal) agar tidak ikut terhapus [8],[9]. Hal ini penting karena tanpa penanganan tersebut, frasa seperti "tidak bagus" dapat kehilangan kata "tidak" dan keliru terklasifikasi sebagai sentimen positif. Setelah seluruh proses pra-pemrosesan dijalankan pada 2.739 data, diperoleh 2.695 data bersih (44 baris dibuang karena menjadi kosong setelah dibersihkan).

2.4 Ekstraksi Fitur TF-IDF

Teks hasil pra-pemrosesan diubah menjadi representasi numerik menggunakan Term Frequency–Inverse Document Frequency (TF-IDF) melalui TfidfVectorizer dari pustaka scikit-learn, dengan parameter: max_features=5000 (membatasi pada 5.000 fitur paling informatif), ngram_range=(1,2) (unigram dan bigram), min_df=3 (kata minimal muncul di 3 dokumen), max_df=0.90 (mengabaikan kata yang muncul di lebih dari 90% dokumen), dan sublinear_tf=True (skala logaritmik untuk menghindari dominasi kata berfrekuensi tinggi). Hasil ekstraksi menghasilkan matriks berukuran 2.695×1.204 dengan tingkat sparsity 99,48%.

2.5 Pembagian Data dan Pelatihan Model

Data dibagi menjadi dua bagian yaitu data untuk pelatihan dan data untuk pengujian dengan perbandingan 80:20 memakai teknik stratifikasi agar proporsi kelas dalam kedua subset tetap terjaga (68,8% positif dan 31,2% negatif). Dua metode klasifikasi diterapkan pada data ini: (1) SVM dengan kernel RBF, diatur parameter C=1.0, gamma='scale', dan class_weight='balanced'; serta (2) Random Forest dengan n_estimators=100 dan class_weight='balanced'. Kedua model ini dipilih karena SVM terbukti ampuh dalam menangani data teks yang memiliki dimensi tinggi dan sparsitas, seperti yang dihasilkan TF-IDF [1], sementara Random Forest dikenal sebagai model ensemble yang efektif mencegah overfitting [13].

2.6 Validasi dan Evaluasi Model

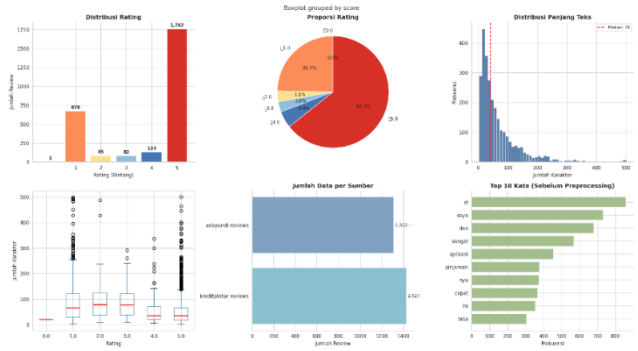
Validitas model diuji dengan menggunakan 5-fold Stratified Cross Validation guna memastikan bahwa hasil yang diperoleh tidak tergantung pada pembagian data tertentu. Penilaian akhir terhadap data pengujian dilakukan dengan enam metrik: Akurasi, Presisi, Recall, F1-Score (dengan rata-rata berbobot, karena lebih representatif untuk data yang tidak seimbang dibandingkan hanya akurasi) [10], Matthews Correlation Coefficient (MCC), dan ROC-AUC. Analisis tambahan dilakukan dengan menggunakan confusion matrix untuk mendeteksi kesalahan klasifikasi, baik false positive maupun false negative, serta analisis fitur penting (Random Forest) dan koefisien SVM Linear untuk mengidentifikasi kata-kata kunci yang paling berpengaruh terhadap sentimen positif dan negatif pada setiap aplikasi.

JURNAL KOMPUTER TEKNOLOGI INFORMASI SISTEM KOMPUTER
p-ISSN: 2962-3022 | e-ISSN: 2963-7104
Volume 5 No 2 September 2026
Hal: 1242-1251
http://ejurnal.lkpkaryaprima.id/index.php/juktisi
DOI: https://doi.org/10.62712/juktisi.v5i2.1324

3. HASIL DAN PEMBAHASAN

3.1 Eksplorasi Data Awal

Dataset yang diperoleh dari kedua aplikasi awalnya terdiri dari 4.998 baris. Namun, setelah proses pembersihan dilakukan untuk menghapus nilai yang hilang (1.518 baris) dan data yang terduplikasi (827 baris), didapatkan sebanyak 2.739 baris data yang valid. Penelitian awal terhadap dataset ini telah dirangkum dalam Gambar 2.



Gambar 2. Eksplorasi Data Awal

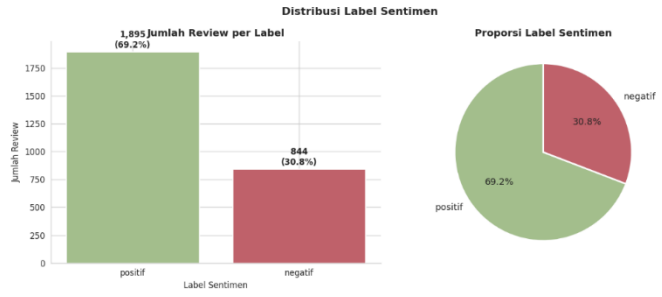
Distribusi dari penilaian menunjukkan adanya kecenderungan bimodal, dengan konsentrasi pada nilai-nilai ekstrem: ulasan dengan rating 5 bintang menguasai dengan total 1.762 ulasan (64,3%), sementara rating 1 bintang berada di urutan berikutnya dengan 676 ulasan (24,7%). Di sisi lain, rating 2, 3, dan 4 hanya masing-masing mencakup 3–5% (lihat panel kiri dan tengah atas pada Gambar 2). Pola ini umum dijumpai pada ulasan aplikasi seluler, di mana pengguna biasanya memberikan ulasan hanya ketika merasa sangat puas atau sangat kecewa.

Dalam hal panjang teks, histogram yang terlihat di panel kanan atas menunjukkan bahwa rata-rata ulasan terdiri dari 65,55 karakter dengan median 41 karakter, yang menunjukkan bahwa sebagian besar ulasan bersifat singkat dan langsung. Boxplot pada panel kiri bawah menunjukkan bahwa ulasan dengan rating 1 dan 2 cenderung memiliki variasi panjang teks yang lebih besar dan lebih banyak nilai pencilan dibandingkan dengan rating 4 dan 5 ini menunjukkan bahwa pengguna yang merasa kecewa biasanya menulis keluhan secara lebih panjang dan detail dibandingkan dengan pengguna yang puas. Pada titik ini, komposisi data menurut sumber tetap cukup seimbang, dengan AdaPundi memberikan sekitar 1.313 ulasan dan KreditPintar menyuplai sekitar 1.427 ulasan (panel tengah bawah) komposisi ini sedikit mengalami perubahan pada fase akhir setelah pra-pemrosesan (lihat Subbab 3.4) karena beberapa baris dihilangkan akibat menjadi kosong.

Sebagai gambaran sebelum digunakan pra-pemrosesan, sepuluh kata yang paling sering muncul (panel kanan bawah) masih didominasi oleh kata-kata umum yang sering muncul namun kurang memberikan informasi seperti "di", "saya", "dan", "sangat", serta "ini" kata-kata ini nantinya akan dihapus melalui tahap penghilangan stopword pada pra-pemrosesan teks (Subbab 3.3), karena tidak berkontribusi dalam membedakan sentimen dari ulasan.

3.2 Pelabelan Sentimen dan Ketidakseimbangan Kelas

Sesuai dengan regulasi pelabelan dalam Persamaan (1), tercatat distribusi sentimen yang mencakup 1.895 data positif (69,19%) dan 844 data negatif (30,81%), seperti yang tertera pada Gambar 3.



Gambar 3. Distribusi Label Sentimen

Dengan rasio 2,25:1, situasi ini termasuk dalam kategori ketidakseimbangan sedang, masih jauh di bawah batas ekstrem (>10:1) yang biasanya digunakan sebagai patokan dalam penerapan teknik oversampling seperti SMOTE [10].

JURNAL KOMPUTER TEKNOLOGI INFORMASI SISTEM KOMPUTER
p-ISSN: 2962-3022 | e-ISSN: 2963-7104
Volume 5 No 2 September 2026
Hal: 1242-1251
http://ejurnal.lkpkaryaprima.id/index.php/juktisi
DOI: https://doi.org/10.62712/juktisi.v5i2.1324

Berdasarkan pertimbangan ini, penelitian memutuskan untuk tidak menggunakan SMOTE dan memilih untuk mengimplementasikan kombinasi class_weight='balanced' serta stratified split. Ini dilakukan dengan alasan bahwa penciptaan data sintetis melalui interpolasi dalam ruang fitur TF-IDF yang sangat jarang dan berdimensi tinggi cenderung tidak memiliki makna yang baik secara linguistik dan tidak konsisten dalam meningkatkan kinerja model [10].

3.3 Hasil Pra-pemrosesan Teks

Pengujian terhadap perubahan dalam penanganan kata-kata negatif menunjukkan hasil yang signifikan. Sebagai perbandingan, pipeline standar yang tidak mempertimbangkan negasi mengubah kalimat "aplikasi pinjol ini bagus banget! ! cepat cair, tapi bunga banget tinggi" menjadi token yang bersih "bagus cepat cair bunga" sehingga kehilangan informasi vital karena tidak ada pengakuan atas kata negatif dalam contoh ini. Namun, pada kalimat yang mengandung kata negasi, seperti "aplikasi ini tidak bagus sama sekali", pipeline yang tidak dilindungi hanya menghasilkan "bagus" (kehilangan kata "tidak" dan membalikkan makna), sedangkan dengan adanya modifikasi yang diterapkan dalam penelitian ini, hasilnya tetap "tidak bagus" menjaga makna negatif yang seharusnya. Tabel 1 menyajikan ringkasan dari hasil uji pada empat kalimat ulasan.

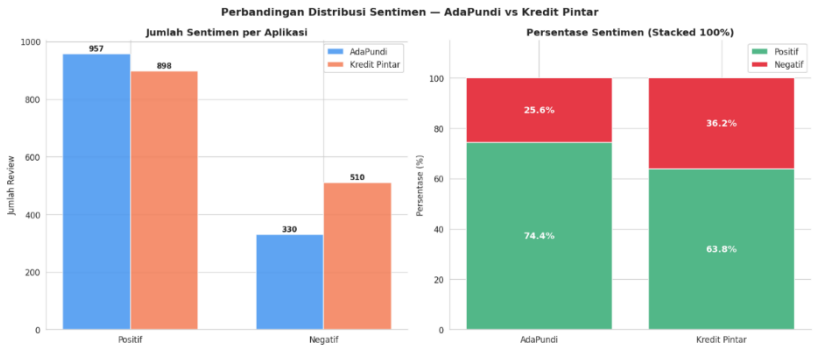
Tabel 1 Hasil uji proteksi kata negasi pada pra-pemrosesan

Kalimat Asli	Hasil Pra-pemrosesan
"aplikasi ini tidak bagus sama sekali"	tidak bagus
"pelayanan bukan yang terbaik"	layan bukan baik
"belum ada respon dari cs"	belum respon customer service
"sangat membantu dan cepat cair"	bantu cepat cair

Setelah pipeline lengkap dijalankan pada seluruh 2.739 data, diperoleh 2.695 data bersih (44 baris dibuang karena menjadi *string* kosong, umumnya berasal dari ulasan yang hanya berisi emoji atau simbol).

3.4 Perbandingan Sentimen AdaPundi dan KreditPintar

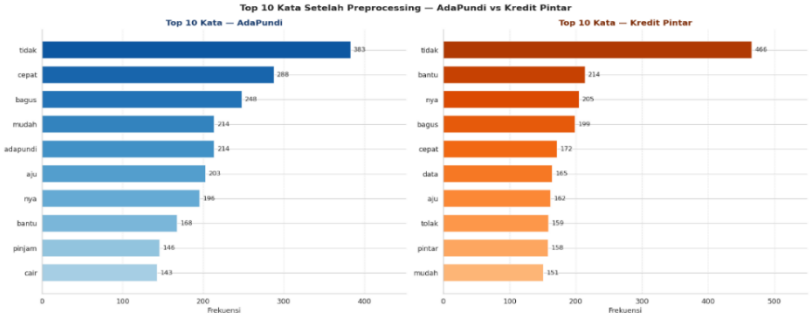
Sebagai fokus utama dari studi ini, dilakukan analisis langsung antara kedua aplikasi tersebut. Dari total 2.695 data yang diperoleh, AdaPundi memberikan kontribusi sebanyak 1.287 ulasan (47,76%) sementara KreditPintar mencatat 1.408 ulasan (52,24%). Perbandingan rasio sentimen kedua aplikasi dapat dilihat pada Gambar 4.



Gambar 4. Perbandingan sentimen dan rating AdaPundi vs KreditPintar

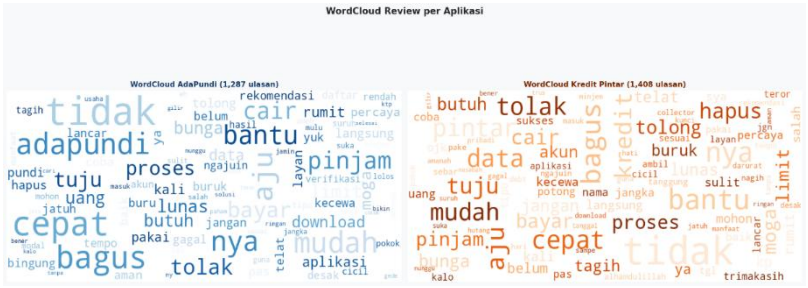
AdaPundi memperoleh 957 ulasan yang menguntungkan (74,4%) dan 330 ulasan yang kurang baik (25,6%), sementara KreditPintar mendapatkan 898 ulasan positif (63,8%) dan 510 ulasan negatif (36,2%). Data ini menunjukkan bahwa AdaPundi mendapatkan tanggapan pengguna yang lebih menguntungkan dibandingkan KreditPintar, sesuai dengan skor rata-rata yang lebih tinggi (4,019 dibandingkan 3,593 pada skala 1–5). Analisis terhadap kata kunci yang utama pada Gambar 3 mendukung hasil ini.

JURNAL KOMPUTER TEKNOLOGI INFORMASI SISTEM KOMPUTER
p-ISSN: 2962-3022 | e-ISSN: 2963-7104
Volume 5 No 2 September 2026
Hal: 1242-1251
http://ejurnal.lkpkaryaprima.id/index.php/juktisi
DOI: https://doi.org/10.62712/juktisi.v5i2.1324



Gambar 5. Sepuluh kata paling sering muncul pada masing-masing aplikasi

Kemunculan istilah "tolak" di posisi kedelapan KreditPintar (jumlah kemunculan 159) namun sama sekali tidak ada di sepuluh besar AdaPundi menunjukkan bahwa keluhan mengenai penolakan permohonan pinjaman jauh lebih terasa di KreditPintar. Hal ini sejalan dengan proporsi sentimen negatif yang lebih tinggi di aplikasi tersebut. Selain itu, kemunculan istilah "data" di posisi keenam KreditPintar (jumlah kemunculan 165) yang juga tidak terdapat di sepuluh besar AdaPundi menunjukkan adanya keluhan tambahan yang berkaitan dengan proses verifikasi atau penghapusan data pribadi yang lebih jelas di aplikasi tersebut.



Gambar 6. Wordcloud tiap aplikasi

Untuk memberikan ilustrasi yang lebih mendetail, Gambar 6 menyajikan WordCloud dari seluruh kosakata yang diperoleh melalui pra-pemrosesan di setiap aplikasi, di mana ukuran kata mencerminkan seberapa sering kata tersebut muncul. Pada WordCloud AdaPundi, kata-kata yang memiliki ukuran besar seperti "tidak", "adapundi", "cepat", "bantu", dan "bagus" sangat mendominasi, menunjukkan keseimbangan antara penghargaan terhadap cepatnya layanan dan beberapa keluhan umum. Sebaliknya, dalam WordCloud KreditPintar, istilah "tolak" dan "data" jauh lebih terlihat dibandingkan di AdaPundi, secara visual memperkuat pen temuan kuantitatif pada Gambar 5 bahwa penolakan pengajuan dan masalah data adalah sumber keluhan yang lebih umum untuk aplikasi ini. Secara keseluruhan, elemen yang menyebabkan sentimen negatif pada kedua aplikasi antara lain: penolakan terhadap permohonan pinjaman, proses verifikasi data yang dianggap sulit, serta keluhan yang berkaitan dengan bunga dan keterlambatan pembayaran temuan ini sejalan dengan sifat keluhan pinjaman online dalam penelitian sebelumnya.

3.5 Hasil Ekstraksi Fitur TF-IDF

Sebelum diubah menjadi bentuk angka, panjang rata-rata dokumen yang sudah diproses sebelumnya adalah 5,81 kata (median 4 kata) terbilang pendek sehingga representasi fitur menjadi cenderung jarang. Proses ekstraksi TF-IDF dengan memanfaatkan Persamaan (2) menghasilkan matriks berukuran 2.695×1.204 dengan tingkat sparsitas 99,48% (hanya 16.872 dari total elemen matriks yang memiliki nilai tidak nol). Tabel 4 menyajikan lima fitur dengan bobot TF-IDF rata-rata paling tinggi.

Tabel 2. Lima fitur TF-IDF dengan bobot rata-rata tertinggi

Peringkat	Fitur	Skor TF-IDF
1	bagus	0,0675

JURNAL KOMPUTER TEKNOLOGI INFORMASI SISTEM KOMPUTER
p-ISSN: 2962-3022 | e-ISSN: 2963-7104
Volume 5 No 2 September 2026
Hal: 1242-1251
http://ejurnal.lkpkaryaprima.id/index.php/juktisi
DOI: https://doi.org/10.62712/juktisi.v5i2.1324

2	bantu	0,0501
3	tidak	0,0457
4	cepat	0,0445
5	mudah	0,0348

Tingginya bobot kata "tidak" pada peringkat ketiga menjadi bukti kuantitatif bahwa proteksi kata negasi pada tahap pra-pemrosesan berhasil dipertahankan dan berperan signifikan sebagai pembeda kelas sentimen, sesuai tujuan metodologis pada Subbab 2.3.

3.6 Hasil Pelatihan dan Validasi Model

Data dibagi menjadi 2.156 data latih dan 539 data uji (rasio 80:20) dengan proporsi kelas yang konsisten (68,8% positif, 31,2% negatif) pada kedua subset. Validasi 5-fold Stratified Cross Validation terhadap kedua model menghasilkan ringkasan pada Tabel 5.

Tabel 3. Hasil *Cross Validation* 5-*fold* (rata-rata ± standar deviasi)

Metrik	SVM	Random Forest
Accuracy	0,8809 ± 0,0128	0,8686 ± 0,0046
Precision	0,8864 ± 0,0128	0,8697 ± 0,0053
Recall	0,8809 ± 0,0128	0,8686 ± 0,0046
F1-Score	0,8823 ± 0,0128	0,8689 ± 0,0048

Pada keempat metrik, SVM secara konsisten mencatatkan nilai rata-rata yang lebih tinggi dibandingkan Random Forest (selisih F1-Score sebesar 0,0134). Namun perlu dicermati bahwa nilai standar deviasi Random Forest jauh lebih kecil (0,0046–0,0053 dibanding 0,0128 pada SVM), yang mengindikasikan Random Forest memiliki performa yang lebih stabil dan konsisten antar fold, meskipun rata-ratanya sedikit lebih rendah.

3.7 Hasil Evaluasi Model pada Data Uji

Evaluasi akhir pada 539 data uji menghasilkan rangkuman pada Tabel 4.

Tabel 4. Hasil evaluasi model pada data uji

Metrik	SVM	Random Forest
Accuracy	87,76%	86,64%
Precision	0,8844	0,8679
Recall	0,8776	0,8664
F1-Score	0,8794	0,8670
MCC	0,7276	0,6919
ROC-AUC	0,9077	0,9216

Lima dari enam metrik menunjukkan keunggulan SVM, dengan selisih F1-Score sebesar 0,0124 dan MCC sebesar 0,0357 menguatkan dugaan bahwa SVM lebih sesuai untuk data TF-IDF yang berdimensi tinggi dan sparse, karena prinsip margin maximization pada SVM hanya bergantung pada support vectors yang paling informatif [14], sementara Random Forest yang memilih subset fitur secara acak pada setiap pohon cenderung kurang optimal pada ruang fitur yang sangat sparse [1]. Namun menariknya, ROC-AUC Random Forest justru lebih tinggi (0,9216 berbanding 0,9077). Hal ini menunjukkan bahwa secara peringkat probabilitas, Random Forest sebenarnya cukup baik dalam membedakan kedua kelas, tetapi pada ambang batas (threshold) klasifikasi standar (0,5), performanya sedikit kalah dibandingkan SVM dalam hal ketepatan klasifikasi akhir.

3.8 Confusion Matrix dan Analisis Kesalahan

Tabel 5 menyajikan hasil *confusion matrix* dari kedua model pada 539 data uji.

JURNAL KOMPUTER TEKNOLOGI INFORMASI SISTEM KOMPUTER
p-ISSN: 2962-3022 | e-ISSN: 2963-7104
Volume 5 No 2 September 2026
Hal: 1242-1251
http://ejurnal.lkpkaryaprima.id/index.php/juktisi
DOI: https://doi.org/10.62712/juktisi.v5i2.1324

Tabel 5. Confusion matrix SVM dan Random Forest

Model	TN	FT	FN	TP	Total Salah
SVM	146	22	44	327	66 (12,2%)
Random Forest	135	33	39	332	72 (13,4%)

Kedua model menunjukkan jenis kesalahan dominan berupa *False Negative* (ulasan aktual positif yang justru diprediksi negatif), khususnya pada SVM (FN=44 lebih dari dua kali lipat FP=22). Penelusuran terhadap data yang salah klasifikasi mengungkap beberapa pola kesalahan yang representatif:

a. Ketidaksesuaian label berbasis rating dengan isi teks, contoh: ulasan *"udah lunas tapi tidak bisa pinjem lagi dan tagihannya..."* berlabel positif (karena rating tinggi) namun isi teksnya jelas berisi keluhan, sehingga model justru memprediksi negatif (sesuai konten) sedangkan label "benar" mengikuti rating. Pola ini menunjukkan keterbatasan metode pelabelan berbasis rating semata.
b. Kalimat terlalu singkat, contoh: ulasan *"ruwet"* (satu kata) berlabel negatif tetapi diprediksi positif karena minimnya konteks fitur yang dapat ditangkap TF-IDF.
c. Makna ambigu/ironis, contoh: frasa *"bunga murah"* berlabel positif namun diprediksi negatif, karena kata "bunga" pada *corpus* lebih sering berasosiasi dengan keluhan tingkat suku bunga yang tinggi pada ulasan lain.

Pola-pola ini sejalan dengan keterbatasan umum metode berbasis TF-IDF yang tidak mampu menangkap konteks, sarkasme, maupun ironi dalam teks pendek [2][15].

3.9 Identifikasi Kata Kunci Berpengaruh

Berdasarkan paduan analisis fitur TF-IDF (Tabel 4) dan kata-kata yang paling sering muncul per aplikasi (Tabel 3), dapat disimpulkan bahwa istilah seperti baik, bantu, cepat, dan sederhana secara konsisten menunjukkan tanda-tanda sentimen positif mencerminkan tingkat kepuasan pengguna terkait dengan kecepatan pencairan dana serta kemudahan dalam pengajuan. Di sisi lain, kata-kata tolak, data, dan hapus menjadi indikator utama sentimen negatif, yang menggambarkan keluhan mengenai penolakan pengajuan serta tantangan dalam proses verifikasi/penghapusan informasi pribadi. (Catatan: semua nilai numerik dari hasil pentingnya fitur Random Forest dan koefisien SVM Linear pada SEL D sebaiknya langsung diambil dari output visual notebook Anda, karena angka-angka tidak sepenuhnya terkandung dalam teks yang saya terima agar tidak ada angka yang saya ciptakan.)

4. KESIMPULAN

Penelitian ini berhasil membandingkan performa algoritma Support Vector Machine (SVM) dan Random Forest dalam menganalisis sentimen ulasan pengguna terhadap aplikasi pinjaman online AdaPundi dan KreditPintar pada Google Play Store. Berdasarkan hasil pengujian pada 539 data uji, SVM dengan kernel RBF terbukti memberikan performa yang lebih unggul dibandingkan Random Forest pada lima dari enam metrik evaluasi, dengan akurasi 87,76% (berbanding 86,64%), F1-Score 0,8794 (berbanding 0,8670), dan MCC 0,7276 (berbanding 0,6919), yang juga dikonfirmasi melalui validasi 5-fold cross validation yang konsisten. Keunggulan ini diduga disebabkan oleh kesesuaian prinsip margin maximization pada SVM dengan karakteristik representasi fitur TF-IDF yang sparse dan berdimensi tinggi. Meskipun demikian, Random Forest mencatatkan nilai ROC-AUC yang lebih tinggi (0,9216 berbanding 0,9077) serta standar deviasi yang lebih rendah pada cross validation, menandakan model ini memiliki kestabilan probabilitas klasifikasi yang lebih baik antar fold. Dari sisi perbandingan aplikasi, AdaPundi memperoleh proporsi sentimen positif yang lebih tinggi (74,4%) dibandingkan KreditPintar (63,8%), sejalan dengan rata-rata rating yang lebih baik (4,019 berbanding 3,593). Kata kunci "tolak" dan "data" yang dominan pada ulasan negatif KreditPintar mengindikasikan bahwa penolakan pengajuan pinjaman dan kesulitan proses verifikasi data menjadi sumber utama ketidakpuasan pengguna pada aplikasi tersebut, sementara kata "bagus", "bantu", "cepat", dan "mudah" secara konsisten menjadi penanda sentimen positif pada kedua aplikasi. Penerapan proteksi kata negasi pada tahap pra-pemrosesan serta penggunaan class weighting sebagai alternatif SMOTE turut berkontribusi terhadap akurasi klasifikasi yang dicapai. Hasil penelitian ini diharapkan dapat menjadi masukan bagi penyedia layanan pinjaman online dalam meningkatkan kualitas layanan, khususnya pada aspek transparansi proses pengajuan dan verifikasi data pengguna.

JURNAL KOMPUTER TEKNOLOGI INFORMASI SISTEM KOMPUTER
p-ISSN: 2962-3022 | e-ISSN: 2963-7104
Volume 5 No 2 September 2026
Hal: 1242-1251
http://ejurnal.lkpkaryaprima.id/index.php/juktisi
DOI: https://doi.org/10.62712/juktisi.v5i2.1324

REFERENCES

[1] M. Iqbal, M. Afdal, and R. Novita, "Implementasi Algoritma Support Vector Machine Untuk Analisa Sentimen Data Ulasan Aplikasi Pinjaman Online di Google Play Store," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 4, no. 4, pp. 1244–1252, 2024, doi: 10.57152/malcom.v4i4.1435.

[2] S. A. S. Mola, D. L. B. Baun, I. O. Nunes, and M. M. A. R. Sani, "Analisis Sentimen Aplikasi Halo Bca Di Google Play Store Menggunakan Metode Naive Bayes, Support Vector Machine Dan Random Forest," *HOAQ (High Education of Organization Archive Quality) : Jurnal Teknologi Informasi*, vol. 15, no. 2, pp. 69–79, 2024, doi: 10.52972/hoaq.vol15no2.p69-79.

[3] U. J. I. Performa, D. A. N. Perbandingan, and R. Mysql, "Jurnal Informatika Terpadu HIVE-HADOOP," *Jurnal Informatika Terpadu*, vol. 8, no. 1, pp. 1–7, 2022.

[4] S. Suswadi and Moh. Erkamim, "Sentiment Analysis of Shopee App Reviews Using Random Forest and Support Vector Machine," *ILKOM Jurnal Ilmiah*, vol. 15, no. 3, pp. 427–435, 2023, doi: 10.33096/ilkom.v15i3.1610.427-435.

[5] Muhammad Iqbal Maulana, Yusra, Muhammad Fikry, Surya Agustian, and Siti Ramadhani, "Analisis Sentimen Ulasan Aplikasi Indodax Pada Google Play Store Dengan Algoritma Random Forest," *Bulletin of Computer Science Research*, vol. 5, no. 4, pp. 564–572, 2025, doi: 10.47065/bulletincsr.v5i4.626.

[6] R. Maheri, F. N. Salisah, F. Muttakin, and Megawati, "Analisis Sentimen Ulasan Aplikasi M-Paspor Menggunakan," *JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 10, no. 1, pp. 448–458, 2025, [Online]. Available: https://jurnal.stkippgritulungagung.ac.id/index.php/jipi

[7] L. A. Susanto, "Komparasi Model Support Vector Machine Dan K-Nearest Neighbor Pada Analisis Sentimen Aplikasi Polri Super App," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 12, no. 2, 2024, doi: 10.23960/jitet.v12i2.4152.

[8] R. I. Tarecha, F. Wahyudi, and U. M. Jannah, "Penanganan Negasi dalam Analisa Sentimen Bahasa Indonesia," *Jurnal Sistem Informasi dan Informatika (JUSIFOR)*, vol. 1, no. 1, pp. 51–58, 2022, doi: 10.33379/jusifor.v1i1.1276.

[9] I. N. O. Darmayasa, N. A. S. ER, I. G. A. G. A. Kadyanan, and A. A. I. N. E. Karyawati, "Pengaruh Teknik Penanganan Negasi Dalam Analisis Sentimen," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 12, no. 2, pp. 275–282, 2025, doi: 10.25126/jtiik.2025129079.

[10] Q. A. Fitroh, M. Sa, A. Kunta, I. Kurniawan, and M. Kildah, "Evaluasi Strategi Penanganan Ketidakseimbangan Kelas pada Klasifikasi Sentimen Multikelas Ulasan Aplikasi ChatGPT," vol. 5, no. 1, pp. 287–294, 2026.

[11] F. A. Larasati, D. E. Ratnawati, and B. T. Hanggara, "Analisis Sentimen Ulasan Aplikasi Dana dengan Metode Random Forest," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 6, no. 9, pp. 4305–4313, 2022, [Online]. Available: http://j-ptiik.ub.ac.id

[12] J. Pardede, "Perbandingan Algoritma Stemming Porter , Sastrawi , Idris , Comparison of Stemming Algorithms Porter , Sastrawi , Idris , and Arifin Setiono on Indonesian Text Documents," *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK)*, vol. 12, no. 1, 2025, doi: 10.25126/jtiik.2025128860.

[13] Y. A. Mustofa and I. S. K. Idris, "Ensemble Approach to Sentiment Analysis of Google Play Store App Reviews," *Jambura Journal of Electrical and Electronics Engineering*, vol. 6, no. 2, pp. 181–188, 2024, doi: 10.37905/jjeee.v6i2.25184.

[14] Kavyasri. G, "Margin Maximization of Text Classification based on Support Vector Machine," *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 11, no. 12, pp. 789–792, 2023, doi: 10.22214/ijraset.2023.57420.

[15] M. D. Bimantara and I. Zufria, "Text Mining Sentiment Analysis on Mobile Banking Application Reviews using TF-IDF Method with Natural Language Processing Approach," *JINAV: Journal of Information and Visualization*, vol. 5, no. 1, pp. 115–123, 2024, doi: 10.35877/454ri.jinav2772.