Idealis: Indonesia Journal Information System
Volume 9, Nomor 2, Juli 2026
ISSN 2684-7280 (*online*)
Halaman 193-203
available online at http://jom.fti.budiluhur.ac.id/index.php/IDEALIS/index

Sentiment Classification of Indonesian PayLater App Reviews Using Random Forest and Logistic Regression

Muhammad Ardi Hermansyah[1*], Muhammad Arifin[2], Pratomo Setiaji[3]

[1,2,3]Sistem Informasi, Fakultas Teknik, Universitas Muria Kudus, Kudus, Indonesia
Email: [1*]muhammadardi931@gmail.com, [2]arifin.m@umk.ac.id, [3]pratomo.setiaji@umk.ac.id
(*: *corresponding author*)

Abstract - *The increasing use of paylater services in Indonesia has generated a large number of user reviews containing meaningful information regarding their experiences. This research focuses on sentiment classification of reviews from three paylater applications using two machine learning approaches: Random Forest and Logistic Regression. Testing was conducted with and without the application of the Synthetic Minority Oversampling Technique (SMOTE) to address class imbalance. Review data was obtained from the Google Play Store and processed through text cleaning, case folding, word normalization, tokenization, stopword removal, and stemming. Feature extraction was performed using the TF-IDF method, followed by splitting the training and testing data using three configurations (80:20, 70:30, and 60:40) to evaluate the consistency of model performance. The evaluation results showed that the 80:20 configuration consistently produced the highest performance. The combination of the Logistic Regression model with SMOTE provided the best results, achieving 88.38% accuracy, 88.37% precision, 88.38% recall, and 88.37% F1-score. Unlike previous studies that generally analyzed a single application without class balancing, this study collected data from three applications simultaneously and empirically compared the effect of SMOTE across various data split configurations, producing a more comprehensive and generalizable benchmark. These findings confirm that SMOTE effectively handles class imbalance, and the combination of TF-IDF, Logistic Regression, and SMOTE is proven to build a highly effective system for classifying the sentiment of paylater application reviews.*

Keywords*: Sentiment Analysis, Paylater, TF-IDF, Logistic Regression, Random Forest, SMOTE.*

Abstrak - Meningkatnya penggunaan layanan *paylater* di Indonesia memunculkan banyak ulasan pengguna yang menyimpan informasi bermakna terkait pengalaman mereka. Riset ini berfokus pada klasifikasi sentimen ulasan tiga aplikasi *paylater* menggunakan dua pendekatan *machine learning*, yakni Random Forest dan Logistic Regression. Pengujian dilakukan dengan dan tanpa penerapan *Synthetic Minority Oversampling Technique* (SMOTE) guna mengatasi ketidakmerataan kelas. Informasi ulasan didapatkan dari Google Play Store, kemudian diproses melalui tahap pembersihan teks, *case folding*, normalisasi kata, tokenisasi, penghapusan *stopword*, dan *stemming*. Ekstraksi fitur dilakukan menggunakan metode TF-IDF, dilanjutkan dengan pembagian data latih dan uji menggunakan tiga konfigurasi (80:20, 70:30, dan 60:40) untuk mengevaluasi konsistensi performa model. Hasil pengujian menunjukkan konfigurasi 80:20 konsisten menghasilkan performa tertinggi. Kombinasi model Logistic Regression dengan SMOTE memberikan hasil paling unggul, mencapai akurasi 88,38%, *precision* 88,37%, *recall* 88,38%, dan *F1-score* 88,37%. Berbeda dengan penelitian terdahulu yang umumnya hanya menganalisis satu aplikasi tanpa penyeimbangan kelas, penelitian ini menghimpun data dari tiga aplikasi sekaligus dan membandingkan secara empiris pengaruh SMOTE pada berbagai konfigurasi pembagian data. Hal ini menghasilkan tolok ukur yang lebih komprehensif dan dapat digeneralisasi. Temuan ini menegaskan bahwa SMOTE efektif menangani ketidakseimbangan kelas, dan kombinasi TF-IDF, Logistic Regression, serta SMOTE terbukti membangun sistem yang sangat efektif untuk mengklasifikasikan sentimen ulasan aplikasi *paylater*.

Kata Kunci: Analisis Sentimen, Paylater, TF-IDF, Logistic Regression, Random Forest, SMOTE.

1. PENDAHULUAN

Kemajuan dalam teknologi keuangan telah mempercepat lahirnya beragam layanan pembayaran elektronik yang mempermudah proses transaksi, salah satunya *paylater*, yaitu metode pembayaran yang memungkinkan pengguna melakukan transaksi terlebih dahulu dan membayarnya pada waktu yang telah ditentukan. Beberapa layanan *paylater* yang banyak digunakan masyarakat Indonesia antara lain Shopee Pay, Kredivo, dan Akulaku[1]. Berbeda dengan layanan pembayaran digital lain seperti dompet elektronik atau transfer bank, *paylater* memiliki karakteristik khusus karena terkait langsung dengan pemberian kredit kepada pengguna, sehingga risiko gagal bayar dan keluhan terhadap proses penagihan menjadi isu yang lebih menonjol[2]. Otoritas Jasa Keuangan mencatat bahwa jumlah kontrak pembiayaan paylater di Indonesia tumbuh rata-rata 144,35% setiap tahun, dari 4,63 juta kontrak pada 2019 menjadi 79,92 juta kontrak pada 2023, sementara data PEFINDO Biro Kredit mencatat jumlah debitur *Buy Now Pay Later* mencapai 17,26 juta orang per Februari 2025[3][4]. Pertumbuhan pengguna yang pesat dan keterlibatan langsung dengan aktivitas kredit inilah yang menjadikan *paylater* objek penelitian yang lebih mendesak untuk dianalisis.

Seiring meningkatnya jumlah pengguna, ulasan mengenai layanan *paylater* pada Google Play Store juga terus bertambah, mulai dari kemudahan penggunaan aplikasi hingga keluhan terkait bunga, penagihan, dan keamanan akun[5]. Banyaknya volume ulasan membuat analisis manual menjadi kurang efisien, sehingga dibutuhkan analisis sentimen untuk mengkategorikan pandangan pengguna ke dalam sentimen positif dan negatif secara otomatis. Penelitian ini memanfaatkan dua algoritma klasifikasi, yakni *Logistic Regression* yang mampu menangani data

Idealis: Indonesia Journal Information System
Volume 9, Nomor 2, Juli 2026
ISSN 2684-7280 (*online*)
Halaman 193-203
available online at http://jom.fti.budiluhur.ac.id/index.php/IDEALIS/index

teks berdimensi tinggi dari hasil ekstraksi fitur TF-IDF secara efektif dan efisien[6], serta *Random Forest* yang menggabungkan banyak pohon keputusan (*desicion tree*) melalui mekanisme voting terbanyak sehingga lebih stabil dan tidak mudah *overfitting*[7]. Untuk mengatasi ketidakseimbangan data pada data ulasan, penelitian ini juga menerapkan *Synthetic Minority Oversampling Technique* (SMOTE) pada data latih[8].

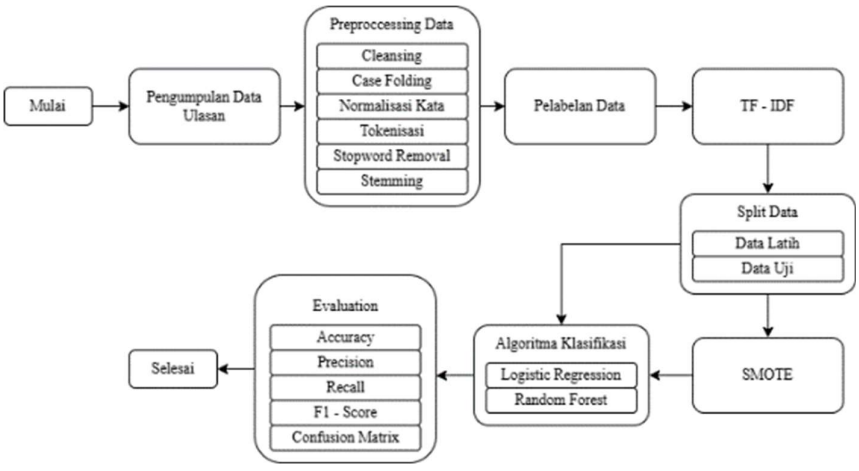
Beberapa penelitian terdahulu telah mengkaji analisis sentimen pada aplikasi finansial sejenis. Penelitian yang dilakukan oleh [9] menganalisis sentimen masyarakat terhadap layanan *paylater* menggunakan *Naive Bayes Classifier* dengan *accuracy* 91%, namun hanya menggunakan satu algoritma tanpa penanganan ketidakseimbangan kelas. Lalu Penelitian oleh [10] meneliti ulasan aplikasi *paylater* yaitu Indodana dengan *accuracy* 87%, sementara [11] menerapkan TF-IDF dan Logistic Regression pada ulasan Shopee dengan *accuracy* 85,11%, namun keduanya tidak menerapkan SMOTE dan hanya berfokus hanya pada satu aplikasi. Berikutnya penelitian tentang perbandingkan *Logistic Regression*, *Naive Bayes*, dan SVM pada ulasan aplikasi *paylater* yaitu Indodana, dengan *Logistic Regression* unggul pada *accuracy* 92,01%, namun tetap berfokus pada satu aplikasi tanpa teknik penyeimbangan data[12]. Penelitian lain pada aplikasi *non-paylater* menerapkan *Random Forest* dan TF – IDF menghasilkan *accuracy* sebesar 68.5%[13]. Penelitian tersebut menunjukkan bahwa performa klasifikasi masih dapat ditingkatkan.

Berdasarkan tinjauan tersebut, terlihat kesenjangan yang belum terjawab yaitu penelitian *paylater* sebelumnya hanya berfokus pada satu aplikasi, sementara penggunaan SMOTE bersama *Logistic Regression* dan *Random Forest* belum diuji khusus pada data ulasan *paylater*. Penelitian ini menutup kesenjangan tersebut dengan mengambil data dari tiga aplikasi *paylater* secara bersamaan (Shopee Pay, Kredivo, dan Akulaku) serta membandingkan performa *Logistic Regression* dan *Random Forest* sebelum dan sesudah penerapan SMOTE, sehingga berkontribusi pada perbandingan empiris yang lebih representatif untuk konteks ulasan *paylater* di Indonesia. Dengan demikian, penelitian ini memberikan kontribusi berupa tolok ukur empiris yang lebih komprehensif dibandingkan penelitian sebelumnya, karena tidak hanya membandingkan dua algoritma secara langsung tetapi juga menguji pengaruh SMOTE pada keduanya menggunakan tiga konfigurasi split data dan data yang lebih representatif dari tiga aplikasi paylater sekaligus.

Berdasarkan penjelasan diatas, penelitian ini dilakukan guna menganalisis sentimen pengguna terhadap aplikasi paylater yang ada di Indonesia dimana datanya diambil dari Google Play Store. Penelitian ini membandingkan kinerja *Logistic Regression* maupun *Random Forest* baik sebelum maupun setelah penerapan SMOTE, guna mengetahui dampak penyeimbangan data terhadap hasil klasifikasi yang dihasilkan. Prosedur penilaian dilakukan menggunakan ukuran accuracy, *precision*, *recall*, dan nilai *F1- Score* untuk memperoleh model dengan performa terbaik dalam mengklasifikasikan sentimen pengguna layanan *paylater*.

2. METODE PENELITIAN

Studi ini dilakukan dengan rangkaian langkah yang terstruktur, dimulai dari pengumpulan data, persiapan awal, pemilahan fitur, pembelajaran model, sampai penilaian hasil klasifikasi. Urutan langkah penelitian tersebut bisa dipahami dalam Gambar 1.



Gambar 1. Alur Penelitian

Gambar 1 menunjukkan alur penelitian diawali dengan pengumpulan data ulasan pengguna aplikasi *paylater* dari Google Play Store. Selanjutnya, data melalui tahap *preprocessing* yaitu *cleansing*, *case folding*, normalisasi kata, tokenisasi, *stopword removal*, serta *stemming* untuk meningkatkan kualitas data. Setelah itu dilakukan pelabelan sentimen dan pembagian data ke data latih serta data uji. Data teks selanjutnya diekstraksi menggunakan

Idealis: Indonesia Journal Information System
Volume 9, Nomor 2, Juli 2026
ISSN 2684-7280 (*online*)
Halaman 193-203
available online at http://jom.fti.budiluhur.ac.id/index.php/IDEALIS/index

metode TF-IDF untuk menghasilkan representasi numerik. Pada tahap berikutnya, teknik SMOTE diterapkan untuk mengatasi ketidakseimbangan kelas sebelum proses klasifikasi menggunakan algoritma *Logistic Regression* serta *Random Forest*. Dan terakhir, kinerja model dievaluasi dengan menggunakan metrik akurasi, *precision*, *recall, F1-score*, serta *confusion matrix*.

2.1. Pengumpulan Data Ulasan

Dalam studi ini, data didapat dengan proses *web scraping* berkaitan dengan pendapat pengguna tentang layanan paylater yang bisa ditemukan di Google Play Store. Aplikasi yang digunakan sebagai sumber data meliputi Shopee Pay, Kredivo, dan Akulaku karena merupakan layanan bayar nanti yang umum dipakai oleh warga Indonesia. Proses pengambilan data dilaksanakan dengan memakai bahasa pemrograman "Python" memanfaatkan pustaka dari *google play scraper* untuk melakukan dengan otomatis. mengumpulkan informasi tentang ulasan pengguna. Data yang dikumpulkan berupa isi ulasan, rating, tanggal ulasan, serta informasi pendukung lainnya yang tersedia pada Google Play Store. Periode data yang dipakai pada studi ini meliputi ulasan yang dipublikasikan mulai 19 Juli 2025 hingga 31 Mei 2026.

Mengingat platform Shopee Pay, Kredivo, dan Akulaku tidak hanya menyediakan fitur Paylater melainkan juga berbagai fitur finansial lainnya (seperti dompet digital, transfer, atau pinjaman umum), maka dilakukan proses filtrasi konten ulasan. Proses penyaringan ini dilakukan dengan memfilter ulasan yang memuat kata kunci spesifik yang berkaitan langsung dengan fitur paylater, antara lain "paylater", "limit", "bunga", serta istilah relevan lainnya. Dari proses pengumpulan dan filtrasi data tersebut, diperoleh sebanyak 7.684 ulasan pengguna yang selanjutnya digunakan sebagai dataset penelitian. Setelah proses pengumpulan data selesai, dataset kemudian melalui tahapan *preprocessing* untuk menghilangkan data yang tidak relevan, duplikasi, dan *noise* sehingga diperoleh data yang lebih bersih dan siap digunakan pada proses analisis sentimen. Hasil dari pengumpulan data terdapat pada Tabel 1.

Tabel 1. *Scrapping Data*

Nama Aplikasi	Username	Rating	Review	Tanggal
Shopee	Pengguna Google	5	sudah bisa cicilan ok	2026-05-30
Kredivo	Pengguna Google	1	bunga ny gede nyett...mahallllll	2026-05-25
Akulaku	Pengguna Google	1	pengajuannya lama, limit kecil, selalu ditolak	2026-05-25

2.2. *Preprocessing* Data

Tahap *preprocessing* data dilaksanakan guna memperbaiki mutu data teks sebelumnya digunakan dalam proses pengelompokan sentimen. Proses ini memiliki tujuan untuk menghilangkan *noise*, menyeragamkan bentuk kata, dan menghasilkan representasi teks yang terstruktur sehingga dapat meningkatkan kinerja algoritma klasifikasi [14]. Dari total 7.684 ulasan yang berhasil dikumpulkan, dilakukan penghapusan 69 data duplikat, 92 ulasan yang terlalu pendek, 430 ulasan dengan rating 3 (netral), serta 424 ulasan yang mengandung noise atau informasi yang ambigu. Setelah proses penyaringan data selesai, diperoleh 6.669 ulasan yang digunakan pada tahap penelitian selanjutnya, yang terdiri dari 3.674 ulasan positif serta 2.995 ulasan negatif.

a. *Cleansing*
Cleansing merupakan tahapan dalam membersihkan data dengan mengeliminasi karakter yang tidak dibutuhkan misalnya tautan, penyebutan, tagar, bilangan, tanda baca, emoji, serta simbol lain yang tidak berpengaruh pada analisis sentimen. Hasil Pembersihan terdapat pada Tabel 2.

Tabel 2. *Cleansing*

Sebelum	Sesudah
masih belum bisa mengajukan paylater karna alasan cacat nama🤭 🤭 tapi sudah 8 tahun lalu kredit macet karna bencana kota palu 2018 kendaraan saya kembalikan ke pihak lising🤭 apakah ada solusi lain? mohon arahannya min tq🙏🙏🙏	masih belum bisa mengajukan paylater karna alasan cacat nama tapi sudah tahun lalu kredit macet karna bencana kota palu kendaraan saya kembalikan ke pihak lising apakah ada solusi lain mohon arahannya min tq

b. *Case Folding*
Ini dilaksanakan dengan mengubah semua huruf dalam berkas menjadi huruf mini agar mencegah adanya variasi dalam penulisan kata yang disebabkan oleh pemakaian huruf besar. Hasil dari Case Folding terdapat pada Tabel 3.

Tabel 3. *Case Folding*

Sebelum	Sesudah
halo ka tolong di ACC dana PAYLATER saya saya peguna lama selalu di tolak saat daftar sopylater tolong di konfirmasi pendaftaran saya	halo ka tolong di acc dana paylater saya saya peguna lama selalu di tolak saat daftar sopylater tolong di konfirmasi pendaftaran saya

Idealis: Indonesia Journal Information System
Volume 9, Nomor 2, Juli 2026
ISSN 2684-7280 (*online*)
Halaman 193-203
available online at http://jom.fti.budiluhur.ac.id/index.php/IDEALIS/index

c. Normalisasi
Normalisasi memiliki tujuan untuk mengonversi kata-kata yang tidak sesuai kaidah, singkatan, serta bahasa yang tidak formal menjadi bentuk yang sesuai standar sesuai dengan kamus normalisasi yang sudah dibuat. Beberapa contoh normalisasi yang dilakukan antara lain "ga", "gak", dan "gk" menjadi "tidak", "dpt" menjadi "dapat", "udh" menjadi "sudah", "bgt" menjadi "banget", serta "apk" menjadi "aplikasi"[15]. Terdapat pada Tabel 4.

Tabel 4. Normalisasi

Sebelum	Sesudah
ga dpt limit sekalinye dpt ga bs di pakai	tidak dapat limit sekalinya dapat tidak bisa di pakai

d. Tokenisasi
Tokenisasi adalah langkah yang membagi teks menjadi sebuah perkumpulan kata sehingga setiap kata dapat ditangani secara terpisah di tahap-tahap analisis selanjutnya. Hasil tokenisasi berupa daftar kata yang digunakan sebagai masukan untuk proses *stopword removal* dan *stemming*. Hasil Tokenisasi dapat ditinjau di Tabel 5.

Tabel 5. Tokenisasi

Sebelum	Sesudah
sudah bisa cicilan ok	['sudah', 'bisa', 'cicilan', 'ok']

e. Stopword *Removal*
Ini dilakukan dengan *delete* kata-kata yang mempunyai frekuensi tinggi tetapi kurang memberi informasi penting terhadap sentimen. Daftar *stopword* yang diterapkan diperoleh dari pustaka Sastrawi serta dikombinasikan dengan berbagai stopword tambahan yang sering muncul pada ulasan pengguna, seperti "yang", "dan", "untuk", "di", "ke", "dengan", "ini", dan "itu" [16]. Contoh proses terdapat di Tabel 6.

Tabel 6. Stopword *Removal*

Sebelum	Sesudah
['aplikasi', 'ini', 'emang', 'mantap', 'mempermudah', 'pembayaran', 'dan', 'cicilan']	['aplikasi', 'emang', 'mantap', 'mempermudah', 'pembayaran', 'cicilan']

f. *Stemming*
Stemming bertujuan mengubah kata berimbuhan ke bentuk dasarnya menggunakan algoritma *stemming* Bahasa Indonesia. Sebagai contoh, kata “membayar”, “dibayar”, dan “pembayaran” diubah menjadi kata dasar “bayar”[16]. Hasil *Stemming* dapat ditinjau di Tabel 7.

Tabel 7. Stemming

Sebelum	Sesudah
['tambahkan', 'limit', 'transaksinya']	['tambah', 'limit', 'transaksi']

2.3. Pelabelan Data

Tahap penandaan dilakukan dengan metode berdasarkan penilaian yang disampaikan oleh pengguna di Google Play Store. Ulasan dengan nilai 1 dan 2 tergolong dalam kategori sentimen negatif, nilai 4 dan 5 dimasukkan ke dalam kategori sentimen positif, sedangkan nilai 3 dipandang sebagai sentimen netral. Karena penelitian ini menggunakan klasifikasi biner, data dengan label netral tidak digunakan dalam proses pemodelan[17]. Untuk meningkatkan kualitas data, dilakukan proses *smart labeling* dengan mendeteksi ulasan yang memiliki ketidaksesuaian antara rating dan isi ulasan. Sebagai contoh, ulasan dengan rating tinggi tetapi mengandung frasa negatif seperti "tidak bisa", "diblokir", atau "penipuan" dianggap sebagai data noise. Sebaliknya, ulasan dengan rating rendah tetapi mengandung frasa positif seperti "sangat bagus" atau "sangat membantu" juga dikategorikan sebagai *noise*.

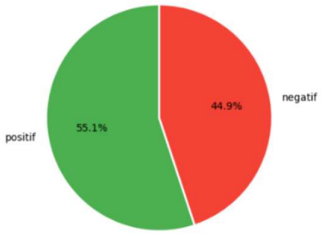
Data yang terdeteksi sebagai *noise* kemudian dihapus dari dataset agar label sentimen yang digunakan lebih konsisten dengan isi ulasan[18]. Dari 7.684 data hasil *preprocessing*, terdeteksi sebanyak 424 data *noise*, 69 data duplikat, 92 ulasan yang terlalu pendek, dan 430 data netral. Setelah proses pembersihan dilaksanakan, diperoleh 6.669 data yang diterapkan pada studi, terdiri dari 3.674 ulasan positif serta 2.995 ulasan negatif.

Tabel 8. Hasil Pelabelan Data

Rating	Ulasan	Hasil
5	aplikasi sangat membantu dan proses cepat	Positif
5	limit tidak muncul dan tidak bisa digunakan	Noise
1	pelayanan buruk dan sering error	Negatif
1	aplikasi sangat bagus dan mudah digunakan	Noise

Idealis: Indonesia Journal Information System
Volume 9, Nomor 2, Juli 2026
ISSN 2684-7280 (*online*)
Halaman 193-203
available online at http://jom.fti.budiluhur.ac.id/index.php/IDEALIS/index

Tabel 8 menampilkan beberapa contoh ulasan beserta label sentimen yang dihasilkan setelah proses pelabelan dan deteksi *noise*.



Gambar 2. Visualisasi Hasil Pelabelan Data

Dari Gambar 2, didapatkan hasil analisis sentimen yang menunjukkan bahwa ulasan positif mencapai 55,1% sedangkan ulasan negatif berada di angka 44,9%.

2.4. Ekstraksi Fitur (TF-IDF)

Ekstraksi Fitur dilakukan dengan memakai metode Frekuensi Istilah-Invers Frekuensi Dokumen. Metode ini mengkonversi data teks menjadi vektor angka dengan memberikan nilai pada setiap kata sesuai dengan seberapa sering kata tersebut muncul dalam dokumen dan seberapa langka kata itu dalam keseluruhan koleksi. Jika sebuah kata muncul sering dalam satu dokumen tetapi jarang terdapat di dokumen lainnya, maka nilai TF-IDF yang diberikan akan semakin besar.[19].

Pada penelitian ini, proses ekstraksi fitur dilaksanakan dengan memanfaatkan *TfidfVectorizer* dari library *Scikit-learn* secara parameter *max_features* sebesar 10.000 untuk membatasi jumlah fitur yang digunakan. Selain itu, digunakan kombinasi *unigram* dan *bigram* (*ngram_range = (1,2)*) sehingga model bisa menangkap informasi dari kata tunggal ataupun berpasangan kata secara berurutan. Parameter *min_df = 2* diterapkan untuk menghilangkan kata yang muncul kurang dari dua kali pada seluruh dokumen, sedangkan *sublinear_tf = True* digunakan untuk melakukan normalisasi frekuensi kata. Selain itu, token negasi yang telah dibentuk pada tahap preprocessing dipertahankan selama proses ekstraksi fitur sehingga frasa seperti tidak_bisa dan tidak_muncul tetap dianggap sebagai satu fitur yang utuh. Hasil ekstraksi menghasilkan matriks TF-IDF berukuran 6.669 dokumen × 10.000 fitur yang selanjutnya digunakan pada proses pemodelan.

2.5. *Split Data*

Prosedur pembagian data dilakukan untuk memisahkan kumpulan data menjadi data yang dipakai untuk latihan serta data yang dipakai untuk pengujian sebelum fase pemodelan dimulai. Data yang digunakan untuk latihan dioptimalkan untuk merancang model klasifikasi, sementara data untuk pengujian bertujuan untuk penilaian kinerja model dalam memprediksi informasi yang belum pernah dikenal sebelumnya. Dalam studi ini, distribusi data diatur dengan proporsi 80:20, berarti 80% data akan dipakai untuk data latih dan 20% sisanya untuk data uji. Proses pembagian ini dilakukan secara acak tetapi tetap mempertahankan proporsi masing-masing kelas sentimen, guna mencegah terjadinya bias pada proses evaluasi model. Selain konfigurasi 80:20, penelitian ini juga menguji dua konfigurasi pembagian data tambahan, yaitu 70:30 dan 60:40, untuk mengevaluasi konsistensi performa model pada jumlah data latih yang berbeda serta memperkuat validitas pemilihan konfigurasi split optimal. [20].

2.6. SMOTE

Kondisi data yang tidak seimbang berpotensi membuat model klasifikasi lebih condong memprediksi kelas yang mendominasi. Guna menangani hal tersebut, diterapkan metode *Synthetic Minority Oversampling Technique* (SMOTE). Metode SMOTE beroperasi dengan mengubah Kumpulan informasi yang tidak merata dengan membuat data buatan di kategori yang lebih sedikit. Tujuannya adalah untuk meningkatkan efektivitas dari teknik klasifikasi yang diterapkan. Akan tetapi, penerapan SMOTE berpotensi menimbulkan *overfitting* akibat data pada kelas minoritas yang mengalami duplikasi [21][22].

2.7. Modeling

Pada penelitian ini, klasifikasi sentimen dilaksanakan dengan memanfaatkan dua algoritma *machine learning,* yakni *Logistic Regression* serta *Random Forest*. *Logistic Regression* merupakan algoritma klasifikasi yang berfungsi memprediksi probabilitas keanggotaan suatu data pada kelas tertentu dengan menggunakan fungsi logistik (sigmoid). Algoritma ini kerap dimanfaatkan dalam analisis sentimen karena sanggup mengatasi data teks yang mempunyai dimensi yang tinggi hasil dari ekstraksi fitur TF-IDF, serta memiliki proses komputasi yang

Idealis: Indonesia Journal Information System
Volume 9, Nomor 2, Juli 2026
ISSN 2684-7280 (*online*)
Halaman 193-203
available online at http://jom.fti.budiluhur.ac.id/index.php/IDEALIS/index

cenderung efisien[23]. Nilai probabilitas suatu data untuk masuk ke dalam kelas positif dapat diperoleh melalui Persamaan (1).

$$P(y = 1|x) = \frac{1}{1+e^{-(\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_n x_n)}} \quad (1)$$

$P(y = 1 \mid x)$ adalah probabilitas data termasuk ke dalam kelas positif, β_0 adalah konstanta (*intercept*), βn merupakan koefisien fitur, dan x_n adalah nilai fitur yang digunakan dalam proses klasifikasi.

Selain *Logistic Regression*, penelitian ini juga menerapkan Random Forest sebagai metode pembanding. *Random Forest* yaitu metode pembelajaran kelompok yang membuat banyak pohon keputusan (*decision tree*) dan menggabungkan hasil prediksi dari setiap pohon melalui proses pemungutan suara mayoritas untuk menetapkan kelas akhir. [24]. Pendekatan ini dapat mengembangkan kestabilan model dan menurunkan kemungkinan overfitting karena keputusan klasifikasi tidak semata-mata berdasarkan satu pohon keputusan. Prediksi akhir *Random Forest* ditentukan menggunakan Persamaan (2).

$$\hat{y} = \text{mode}(T_1, T_2, \ldots, T_n) \quad (2)$$

$\hat{y}$ merupakan hasil prediksi akhir, $T_1, T_2, \ldots, T_n$ adalah output prediksi dari setiap pohon keputusan, sementara "modus" menampilkan kategori yang paling sering dipilih oleh semua pohon. Kedua algoritma tersebut kemudian dilatih menggunakan fitur TF-IDF dan dievaluasi performanya baik sebelum maupun sesudah penerapan SMOTE untuk menentukan model terbaik dalam klasifikasi sentimen ulasan pengguna aplikasi *paylater* di Indonesia.

2.8. *Evaluation*

Kinerja model dievaluasi menggunakan *confusion matrix*, yaitu tabel yang menyatakan jumlah prediksi yang benar maupun salah yang dihasilkan dari model klasifikasi. Berdasarkan *confusion matrix*, diperoleh beberapa metrik evaluasi, Nilai *accuracy* dihitung menggunakan Persamaan (3), *precision* dihitung menggunakan Persamaan (4), sedangkan *recall* maupun *F1-Score* dihitung dengan memanfaatkan Persamaan (5) dan Persamaan (6). Indikator ini dimanfaatkan untuk menilai sejauh mana model mampu secara akurat mengkategorikan perasaan positif maupun negatif. [7][25].

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

$$Precision = \frac{TP}{TP+FP} \quad (4)$$

$$Recall = \frac{TP}{TP+FN} \quad (5)$$

$$F1 - Score = \frac{2\times Precision\times Recall}{Precision+Recall} \quad (6)$$

Keterangan :
TP (*True Positive*) : data positif yang diprediksi positif
TN (*True Negative*) : data negatif yang diprediksi negatif
FP (*False Positive*) : data negatif yang diprediksi positif
FN (*False Negative*) : data positif yang diprediksi negatif

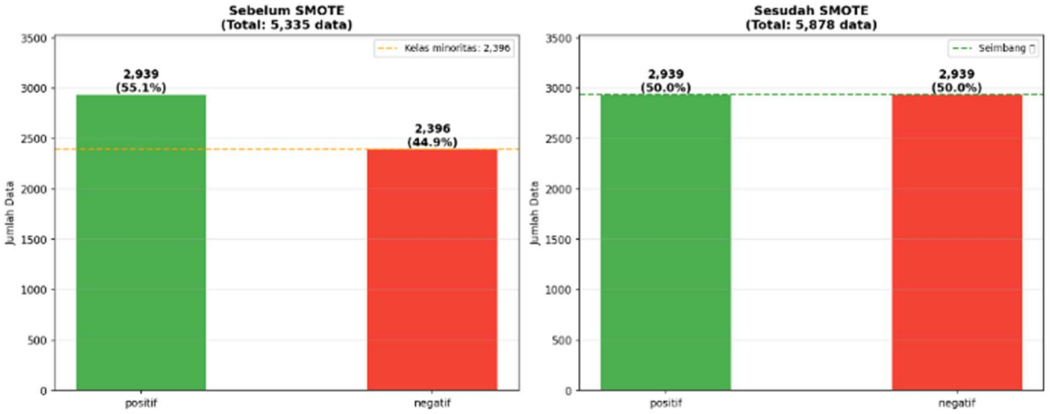
3. HASIL DAN PEMBAHASAN

3.1. Dataset

Pada studi ini, data ulasan pengguna Shopee Pay, Kredivo, dan Akulaku didapat dengan proses mengambil data dari Google Play Store yang memiliki total 7.684 ulasan. Sebelum dilakukan pemodelan, data melalui serangkaian tahapan *preprocessing* untuk meningkatkan kualitas dataset. Pada tahap ini, sebanyak 69 ulasan duplikat, 92 ulasan yang terlalu pendek, 430 ulasan netral (rating 3), serta 424 ulasan yang teridentifikasi sebagai data ambigu atau *noise* dihapus dari dataset. Setelah tahap pembersihan data dilaksanakan, berhasil dikumpulkan 6. 669 ulasan yang menjadi subjek penelitian, yang terdiri dari 3. 674 ulasan positif (55,1%) serta 2. 995 ulasan negatif (44,9%). Meskipun perbedaan jumlah kedua kelas tidak terlalu besar, distribusi data masih menunjukkan dominasi kelas positif. Oleh sebab itu, teknik *Synthetic Minority Oversampling Technique* (SMOTE) digunakan terhadap data pelatihan guna menyeimbangkan distribusi data dengan cara menambahkan contoh sintetis ke dalam kelas yang lebih sedikit.

Penerapan SMOTE bertujuan agar model klasifikasi dapat mempelajari karakteristik kedua kelas secara lebih seimbang serta mengurangi kecenderungan prediksi terhadap kelas mayoritas. Perbandingan distribusi data sebelum maupun sesudah penerapan SMOTE bisa ditinjau pada Gambar 3.

Idealis: Indonesia Journal Information System
Volume 9, Nomor 2, Juli 2026
ISSN 2684-7280 (*online*)
Halaman 193-203
available online at http://jom.fti.budiluhur.ac.id/index.php/IDEALIS/index



Gambar 3. Hasil Penerapan SMOTE

3.2. Evaluasi Perbandingan Model Klasifikasi

Dalam studi ini dilaksanakan evaluasi performa dari kedua algoritma klasifikasi, yaitu algoritma *Logistic Regression* dan juga *Random Forest,* baik sebelum maupun sesudah dilakukan SMOTE. Evaluasi dilakukan untuk mengetahui pengaruh penyeimbangan data terhadap kemampuan model untuk mengelompokkan perasaan dari ulasan pengguna tentang aplikasi paylater. Kinerja setiap model diperiksa berdasarkan akurasi, ketepatan, penarikan kembali, dan F1-Score. Rangkuman hasil pengujian dan perbandingan kinerja seluruh model ditampilkan di Tabel 9.

Tabel 9. Perbandingan Hasil Klasifikasi

Model	*Accuracy*	*Precision*	*Recall*	*F1-Score*
Logistic Regression	87.93%	87.94%	87.93%	87.90%
Logistic Regression + SMOTE	88.38%	88.37%	88.38%	88.37%
Random Forest	87.03%	87.02%	87.03%	87.03%
Random Forest + SMOTE	87.48%	87.49%	87.48%	87.48%

Berdasarkan hasil pengujian, *Logistic Regression* + SMOTE menghasilkan kinerja terbaik dengan akurasi 88,38% serta F1-Score 88,37%. Penerapan SMOTE terbukti mampu meningkatkan kinerja *Logistic Regression* maupun *Random Forest* melalui penyeimbangan distribusi kelas. Meskipun seluruh model menunjukkan hasil yang baik, kombinasi *Logistic Regression* dan SMOTE menjadi pendekatan paling efektif dalam penelitian ini karena memperoleh nilai akurasi*, Precision, Recall*, serta *F1-Score* tertinggi dibandingkan model lain.

Tabel 10. Perbandingan Hasil Klasifikasi Pada Tiga Konfigurasi Split Data

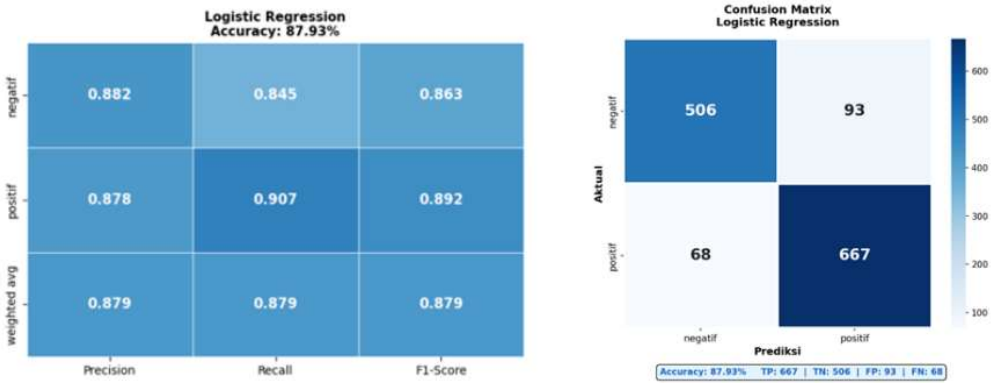
Model	Metrik	80:20	70:30	60:40
Random Forest	*Accuracy*	87,03%	86,61%	86,02%
	Precision	87,02%	86,60%	86,01%
	Recall	87,03%	86,61%	86,02%
	F1-Score	87,03%	86,60%	86,01%
Random Forest + SMOTE	*Accuracy*	87,48%	86,66%	86,09%
	Precision	87,49%	86,66%	86,10%
	Recall	87,48%	86,66%	86,09%
	F1-Score	87,48%	86,66%	86,10%
Logistic Regression	*Accuracy*	87,93%	87,91%	87,18%
	Precision	87,94%	87,91%	87,17%
	Recall	87,93%	87,91%	87,18%
	F1-Score	87,90%	87,88%	87,16%
Logistic Regression + SMOTE	*Accuracy*	**88,38%**	**87,86%**	**87,37%**
	Precision	**88,37%**	**87,85%**	**87,36%**
	Recall	**88,38%**	**87,86%**	**87,37%**
	F1-Score	**88,37%**	**87,85%**	**87,37%**

Tabel 10 menampilkan perbandingan performa keempat model pada tiga konfigurasi split data. Secara keseluruhan, konfigurasi 80:20 menghasilkan performa tertinggi pada semua model dibandingkan 70:30 dan 60:40. Penurunan performa yang terjadi seiring berkurangnya proporsi data latih menunjukkan bahwa jumlah data

Idealis: Indonesia Journal Information System
Volume 9, Nomor 2, Juli 2026
ISSN 2684-7280 (*online*)
Halaman 193-203
available online at http://jom.fti.budiluhur.ac.id/index.php/IDEALIS/index

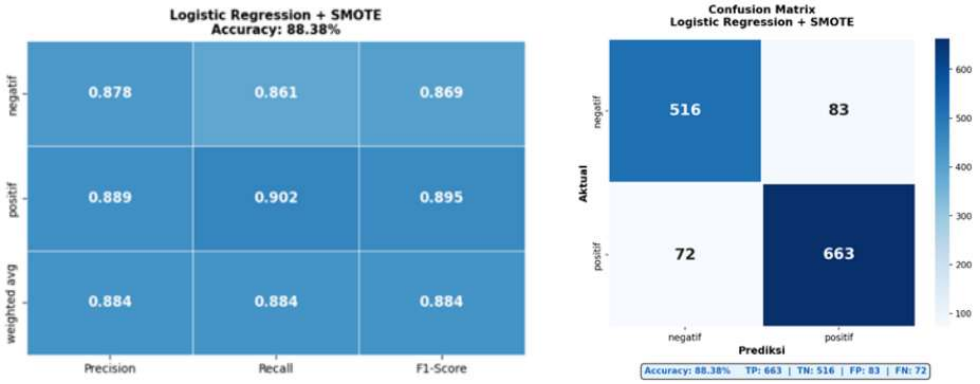
latih yang lebih besar memberikan kontribusi positif terhadap kemampuan model dalam mengenali pola sentimen. Oleh karena itu, konfigurasi 80:20 ditetapkan sebagai konfigurasi optimal dalam penelitian ini.

Model Logistic Regression memperoleh *accuracy* sebesar 87,93%, dengan *precision* 87,94%, *recall* 87,93%, serta *F1-Score* 87,90%. Hasil tersebut menyatakan bahwa model bisa melaksanakan klasifikasi sentimen dengan baik serta menghasilkan prediksi yang relatif konsisten pada kedua kelas. Adapun hasil *Classification Report* dan *Confusion Matrix* terdapat pada Gambar 4 berikut.



Gambar 4. *Classification Report* dan *Confusion Matrix Logistic Regression*

Berdasarkan Gambar 5, hasil pengujian memaparkan bahwa penerapan SMOTE mampu meningkatkan kinerja *Logistic Regression* dengan akurasi sebesar 88,38%, lebih tinggi dibandingkan model tanpa SMOTE (Gambar 4) yang memperoleh *accuracy* 87,93%. Peningkatan juga terjadi pada value *precision*, *recall*, serta *F1-Score*, menyatakan bahwa model bisa melaksanakan klasifikasi sentimen dengan lebih baik. Hasil *Classification Report* dan *Confusion Matrix* terdapat pada Gambar 5.

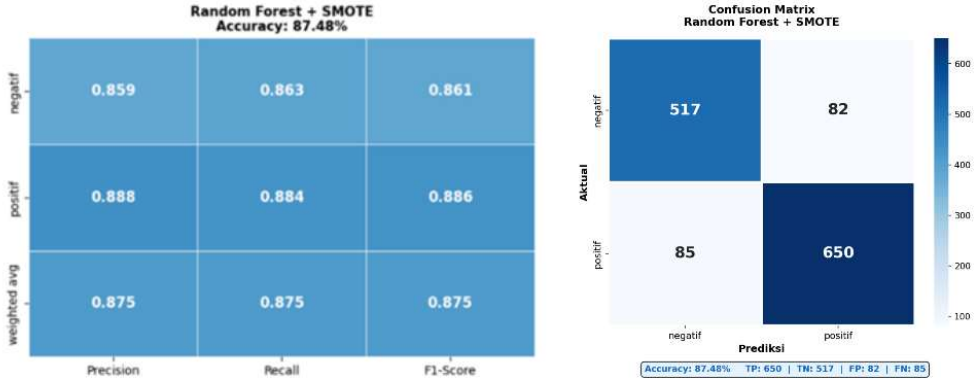


Gambar 5. *Classification Report* dan *Confusion Matrix Logistic Regression* + SMOTE

Gambar 6. *Classification Report* dan *Confusion Matrix Random Forest*

Idealis: Indonesia Journal Information System
Volume 9, Nomor 2, Juli 2026
ISSN 2684-7280 (*online*)
Halaman 193-203
available online at http://jom.fti.budiluhur.ac.id/index.php/IDEALIS/index

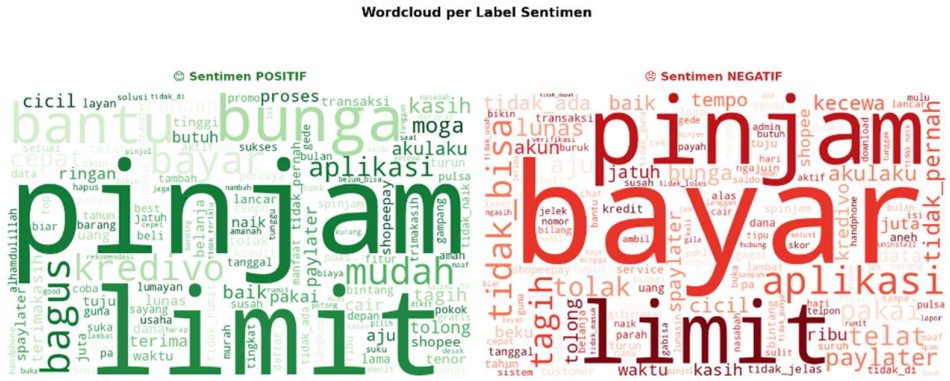
Berdasarkan Gambar 6, model Random Forest memperoleh *accuracy* sebesar 87,03%, dengan *precision* 87,02%, *recall* 87,03%, serta *F1-Score* 87,03%. Nilai tersebut menyatakan bahwa model bisa melaksanakan klasifikasi sentimen dengan baik serta menghasilkan prediksi yang relatif konsisten pada kedua kelas. Hasil *Classification Report* dan *Confusion Matrix* terdapat pada Gambar 6. Berdasarkan Gambar 7, hasil uji model *Random Forest* + SMOTE menghasilkan tingkat *accuracy* sebesar 87,48%, dengan *precision* 87,49%, *recall* 87,48%, serta *F1-Score* 87,48%. Nilai ini menunjukkan bahwa penerapan SMOTE memberikan peningkatan performa dibandingkan *Random Forest* tanpa SMOTE. Adapun hasil *Classification Report* dan *Confusion Matrix* ditunjukkan dalam Gambar 7.



Gambar 7. *Classification Report* dan *Confusion Matrix Random Forest* + SMOTE

3.3. Visualisasi Data

Pada segmen penggambaran data ini, analisis dilakukan dengan menggunakan wordcloud untuk melihat kata-kata yang paling banyak timbul dalam setiap jenis emosi. Gambar tersebut dimaksudkan untuk memberikan gambaran awal terkait persepsi pengguna terhadap layanan *paylater* berdasarkan tingkat frekuensi kemunculan kata dalam ulasan. Ukuran kata yang terlihat pada awan kata mencerminkan seberapa sering kata tersebut muncul, dengan kata yang besar menegaskan frekuensi yang tinggi, semakin tinggi pula frekuensi kemunculannya di dalam dataset.



Gambar 8. *Wordcloud* Positif dan Negatif

Berdasarkan Gambar 8, sentimen yang baik lebih banyak dipengaruhi oleh kata-kata misalnya pinjaman, batas, bunga, gampang, serta bagus, yang menegaskan bahwa para pengguna menganggap layanan paylater ini positif. memberikan kemudahan dan manfaat dalam memenuhi kebutuhan transaksi. Sedangkan, pada sentimen negatif, kata-kata yang dominan diantaranya *bayar*, *tolak*, *telat*, *kecewa*, dan *limit*, yang mengindikasikan bahwa keluhan pengguna banyak berkaitan dengan masalah pembayaran, penolakan pengajuan, serta kendala penggunaan layanan. Hasil visualisasi ini menunjukkan adanya perbedaan karakteristik kata pada masing-masing sentimen dan memberikan gambaran umum mengenai faktor-faktor yang memengaruhi persepsi pengguna pada layanan *paylater*.

3.4. Diskusi

Hasil perbandingan tiga konfigurasi split data menunjukkan bahwa split 80:20 secara konsisten menghasilkan performa tertinggi dibandingkan 70:30 dan 60:40 pada semua model. *Logistic Regression* + SMOTE mencapai

akurasi tertinggi sebesar 88,38% pada split 80:20, menurun menjadi 87,86% pada split 70:30, dan 87,37% pada split 60:40. Penurunan ini sejalan dengan prinsip umum dalam *machine learning* bahwa semakin banyak data latih yang tersedia, semakin baik kemampuan model dalam mempelajari pola klasifikasi. Temuan ini menegaskan bahwa konfigurasi 80:20 merupakan pilihan yang optimal untuk dataset penelitian ini.

Hal menarik ditemukan pada split 70:30, di mana *Logistic Regression* tanpa SMOTE (87,91%) justru sedikit mengungguli *Logistic Regression* + SMOTE (87,86%). Fenomena ini kemungkinan disebabkan oleh berkurangnya jumlah data latih pada split 70:30 sehingga data sintetis yang dihasilkan SMOTE menjadi kurang representatif, dan justru menambahkan sedikit *noise* yang menurunkan performa model. Hal ini menunjukkan bahwa efektivitas SMOTE bergantung tidak hanya pada tingkat ketidakseimbangan kelas, tetapi juga pada ukuran data latih yang tersedia.

Secara keseluruhan, *Logistic Regression* secara konsisten mengungguli *Random Forest* pada seluruh konfigurasi split. Hal ini dapat dijelaskan secara teoritis: representasi TF-IDF menghasilkan matriks fitur yang sangat berdimensi tinggi dan *sparse*, di mana *Logistic Regression* bekerja secara linier dan efisien karena hanya perlu mempelajari satu set bobot untuk memisahkan kelas. Sebaliknya, *Random Forest* yang membangun banyak pohon keputusan pada data *sparse* cenderung kurang optimal karena sebagian besar fitur bernilai nol di tiap node. Temuan ini sejalan dengan penelitian oleh[11] yang juga menemukan *Logistic Regression* dengan TF-IDF efektif pada klasifikasi sentimen ulasan aplikasi dengan akurasi 85,11%, meskipun hasil penelitian ini lebih tinggi karena penerapan *smart labeling* yang membersihkan *noise* label secara lebih sistematis. Di sisi lain, hasil ini sedikit di bawah penelitian oleh[12] yang memperoleh akurasi 92,01% pada ulasan satu aplikasi (Indodana), yang wajar mengingat penelitian ini menggunakan data dari tiga aplikasi berbeda sehingga variasi ulasan lebih tinggi dan tugas klasifikasi lebih menantang namun hasilnya lebih dapat digeneralisasi.

4. KESIMPULAN

Dari hasil pembahasan yang sudah dilaksanakan, teknik Logistic Regression dan Random Forest dapat diandalkan untuk mengklasifikasikan sentimen dalam *review* pengguna aplikasi paylater di Indonesia yang berasal dari Shopee Pay, Kredivo, dan Akulaku. Hasil pengujian mengindikasikan bahwa penerapan SMOTE meningkatkan kinerja kedua teknik tersebut dibandingkan dengan model yang tidak menerapkan penyeimbangan data, yang terdiri dari peningkatan nilai akurasi, presisi, *recall*, serta F1-*score*. Dari keempat skenario yang dievaluasi, model *Logistic Regression* + SMOTE mencapai hasil paling optimal dengan tingkat akurasi 88,38%, presisi 88,37%, recall 88,38%, dan F1-score 88,37%. Dibandingkan dengan penelitian terdahulu yang hanya fokus pada satu aplikasi, penelitian ini memberikan kontribusi tambahan berupa perbandingan empiris antara dua algoritma dengan dan tanpa SMOTE pada data yang dikumpulkan dari tiga aplikasi *paylater* sekaligus.

Dari perspektif praktis, dominasi kata-kata seperti limit, bunga, dan kecewa pada sentimen negatif mengindikasikan bahwa aspek limit kredit, transparansi bunga, dan kecepatan respons layanan pelanggan merupakan prioritas utama yang perlu diperbaiki oleh penyedia layanan *paylater* untuk meningkatkan kepuasan pengguna. Selain itu, analisis visual data mengungkapkan bahwa kata-kata terkait kemudahan penggunaan, limit, dan bunga mendominasi sentimen positif, sedangkan sentimen negatif lebih banyak dipengaruhi oleh permasalahan pembayaran, keterlambatan tagihan, dan penolakan pengajuan pinjaman.

Dengan demikian, kombinasi TF-IDF, *Logistic Regression*, dan SMOTE bisa diterapkan sebagai pendekatan yang baik untuk mengklasifikasikan sentimen ulasan pengguna *paylater* pada dataset penelitian ini. Penelitian ini memiliki beberapa keterbatasan yang perlu diakui. Data hanya dikumpulkan dari Google Play Store sehingga tidak merepresentasikan pengguna platform iOS atau kanal ulasan lainnya. Selain itu, pelabelan sentimen dilakukan secara otomatis berbasis rating pengguna sehingga berpotensi mengandung noise pada ulasan yang isi teksnya tidak konsisten dengan rating yang diberikan, meskipun telah dilakukan proses smart labeling untuk meminimalkan hal tersebut.

Ketiga, penelitian ini hanya menguji dua algoritma klasifikasi dengan fitur TF-IDF, sehingga kemungkinan adanya algoritma atau metode ekstraksi fitur yang lebih optimal belum dieksplorasi. Penelitian selanjutnya dapat mempertimbangkan penggunaan dataset yang lebih besar, algoritma klasifikasi lain, maupun metode ekstraksi fitur yang berbeda untuk mendapatkan hasil yang lebih maksimum.

DAFTAR PUSTAKA

[1] A. Koswara, E. S. Soegoto, R. Wahdiniwaty, M. Bachtiar, and I. D. Sumitra, "Persaingan PayLater di Indonesia: Analisis pasar dan peramalan minat konsumen berbasis data science," *J. Salingka Nagari*, vol. 04, no. 2, pp. 254–270, 2025.

[2] F. J. Wahidna and P. Nerisafitra, "Analisis sentimen pengguna sistem Pay Later menggunakan Support Vector Machine metode pembobotan lexicon," *JINACS (Journal Informatics Comput. Sci.)*, vol. 04, no. 03, pp. 334–343, 2023, doi: 10.26740/jinacs.v4n03.p334-343.

[3] Otoritas Jasa Keuangan, "Roadmap Pengembangan dan Penguatan Perusahaan Pembiayaan 2024–2028," Jakarta, 2024. Accessed: Jun. 28, 2026. [Online]. Available: https://www.ojk.go.id/id/berita-dan-kegiatan/info-

terkini/Documents/Pages/Roadmap-Pengembangan-dan-Penguatan-Perusahaan-Pembiayaan-2024-2028/Roadmap Pengembangan dan Penguatan Perusahaan Pembiayaan 2024-2028_Final.pdf

[4] A. I. Rahmana, "Penyaluran pembiayaan buy now pay later tembus Rp 36,24 triliun per Februari 2025." Accessed: Jun. 28, 2026. [Online]. Available: https://keuangan.kontan.co.id/news/penyaluran-pembiayaan-buy-now-pay-later-tembus-rp-3624-triliun-per-februari-2025

[5] Y. Tandiarny, Afifah, and Y. Saharaeni, "Analisis tingkat kepuasan pengguna Shopee Paylater menggunakan metode User Experience Questionare," *J. Ilmu Komput. KHARISMA Tech*, vol. 29, no. 01, pp. 1–14, 2025, doi: 10.55645/kharismatech.v20i1.505.

[6] I. Ristiana and Mutmainah, "Analisis sentimen bencana banjir Sumatera menggunakan TF-IDF dan Logistic Regression," *J. Sist. Inf. DAN Teknol. (S I N T E K)*, vol. 6, no. 1, pp. 57–65, 2021, doi: 10.56995/sintek.v6i1.239.

[7] T. F. Basar, D. E. Ratnawati, and I. Arwani, "Analisis sentimen pengguna Twitter terhadap pembayaran cashless menggunakan Shopeepay dengan algoritma Random Forest," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 6, no. 3, pp. 1426–1433, 2022.

[8] S. G. Wijaya and Suharyadi, "Analisis sentimen pengguna Twitter terhadap kebijakan royalti restoran dan kafe dengan Multinomial Naive Bayes," *Idealis Indones. J. Inf. Syst.*, vol. 9, no. 1, pp. 49–58, 2026, doi: 10.36080/idealis.v9i1.3698.

[9] A. Safira and F. N. Hasan, "Analisis sentimen masyarakat terhadap Paylater menggunakan metode Naive Bayes Classifier," *Zo. J. Sist. Inf.*, vol. 5, no. 1, pp. 59–70, 2023, doi: 10.31849/zn.v5i1.12856.

[10] R. Rizaldi and R. Aryanti, "Analisis sentimen pengguna terhadap aplikasi Indodana di Google Play Store menggunakan metode Naive Bayes Classifier," *J. Informatics Manag. Inf. Technol.*, vol. 3, no. 4, pp. 98–105, 2024, doi: 10.47065/jimat.v4i3.400.

[11] S. F. Kadir and A. Fairuzabadi, "Analisis sentimen ulasan Shopee di Google Play dengan TF-IDF dan Logistic Regression," *J. Artif. Intell. Digit. Bus.*, vol. 4, no. 2, pp. 7940–7945, 2025, doi: 10.31004/riggs.v4i2.2850.

[12] M. Sahrul and Afiyati, "Sentimen analisis pada ulasan aplikasi Indodana di Google Play Store menggunakan algoritma Logistic Regression, Naïve Bayes dan Support Vector Machine," *J. Nas. Teknol. Komput.*, vol. 5, no. 3, pp. 136–147, 2025, doi: 10.61306/jnastek.v5i3.194.

[13] M. F. Rahmatullah, P. L. Lokapitasari Belluano, and H. Darwis, "Analisis sentimen review aplikasi di Google Play Store menggunakan Random Forest," *LINIER Lit. Inform. dan Komput.*, vol. 2, no. 3, pp. 380–389, 2025, doi: 10.33096/linier.v2i3.3149.

[14] D. Rifaldi, A. Fadhil, and Herman, "Teknik preprocessing pada Text Mining menggunakan Tweet 'Mental Health,'" *Decod. J. Pendidik. Teknol. Inf. ISSN*, vol. 3, no. 2, pp. 161–171, 2023, doi: 10.51454/decode.v3i2.131.

[15] N. Garg and K. Sharma, "Text pre-processing of multilingual for sentiment analysis based on social network data," *Int. J. Electr. Comput. Eng.*, vol. 12, no. 1, pp. 776–784, 2022, doi: 10.11591/ijece.v12i1.pp776-784.

[16] R. Kalaivani and R. Marivendan, "The effect of stop word removal and stemming in datapreprocessing," *Ann. Rom. Soc. Cell Biol.*, vol. 25, no. 6, pp. 739–746, 2021.

[17] K. H, Y. Desnelita, and D. Oktarina, "Analisis sentimen terhadap ulasan aplikasi IKD di Play Store menggunakan Random Forest," *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 14, no. 3, pp. 171–180, 2025, doi: 10.22146/jnteti.v14i3.20473.

[18] H. Wang, B. Liu, Y. Yang, and T. Li, "Learning with noisy labels for sentence-level sentiment classification," *Inf. Process. Manag.*, vol. 60, no. 2, pp. 6286–6292, 2023, doi: 10.18653/v1/D19-1655.

[19] O. I. Gifari, M. Adha, I. R. Hendrawan, and F. F. S. Durrand, "Analisis sentimen review film menggunakan TF-IDF dan Support Vector Machine," *JIFOTECH (Journal Inf. Technol.)*, vol. 2, no. 1, pp. 36–40, 2022, doi: 10.46229/jifotech.v2i1.330.

[20] N. Nurdiansyah, F. S. Febriyan, Z. G. D. Amanta, D. A. Saputra, and W. M. Baihaqi, "Analisis kesehatan mental untuk mencegah gangguan mental pada mahasiswa menggunakan algoritma K-Nearest Neighbor (K-NN) dan Random Forest," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 5, no. 1, pp. 1–10, 2025, doi: 10.57152/malcom.v5i1.1537.

[21] C. Cahyaningtyas, Y. Nataliani, and I. R. Widiasari, "Analisis sentimen pada rating aplikasi Shopee menggunakan metode Decision Tree berbasis SMOTE," *AITI J. Teknol. Inf.*, vol. 18, no. 2, pp. 173–184, 2021, doi: 10.24246/aiti.v18i2.173-184.

[22] A. Fathkudin, F. A. Artanto, N. A. Safli, and D. Wibowo, "Decision Tree berbasis SMOTE dalam analisis sentimen penggunaan artificial intelligence untuk skripsi," *Remik Ris. dan E-Jurnal Manaj. Inform. Komput.*, vol. 8, no. 2, pp. 494–505, 2024, doi: 10.33395/remik.v8i2.13531.

[23] A. Khaidar, "Analisis sentimen di Instagram terhadap menteri keuangan Purbaya Yudhi Sadewa menggunakan metode Logistic Regression," *JITET (Jurnal Inform. dan Tek. Elektro Ter.)*, vol. 13, no. 3, pp. 1674–1683, 2025, doi: 10.23960/jitet.v13i3S1.8002.

[24] N. A. Hapsari and A. D. Indriyanti, "Analisis sentimen pada aplikasi dompet digital menggunakan algoritma Random Forest," *J. Emerg. Inf. Syst. Bus. Intell.*, vol. 4, no. 3, pp. 186–192, 2023, doi: 10.26740/jeisbi.v4i3.55696.

[25] N. Hidayah and Dodiman, "Implementasi algoritma Multinomial Naïve Bayes, TF-IDF dan Confusion Matrix dalam pengklasifikasian saran monitoring dan evaluasi mahasiswa terhadap dosen Teknik Informatika Universitas Dayanu Ikhsanuddin," *J. Akad. Pendidik. Mat.*, vol. 10, no. 1, pp. 8–15, 2024, doi: 10.55340/japm.v10i1.1491.