

© Jurnal Nasional Teknik Elektro dan Teknologi Informasi
Karya ini berada di bawah Lisensi Creative Commons Atribusi-BerbagiSerupa 4.0 Internasional
Terjemahan artikel 10.22146/jnteti.v14i3.20473

Analisis Sentimen Terhadap Ulasan Aplikasi IKD di Play Store Menggunakan *Random Forest*

Kelvin H.¹, Erlin², Yenny Desnelita¹, Dwi Oktarina¹

¹ Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Institut Bisnis dan Teknologi Pelita Indonesia, Pekanbaru, Riau 28127, Indonesia

² Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Institut Bisnis dan Teknologi Pelita Indonesia, Pekanbaru, Riau 28127, Indonesia

[Diserahkan: 26 Maret 2025, Direvisi: 13 Juni 2025, Diterima: 24 Juli 2025]

Penulis Korespondensi: Kelvin H. (email: kelvin.h@student.pelitaindonesia.ac.id)

INTISARI — Pertumbuhan pesat aplikasi digital dalam layanan administrasi kependudukan meningkatkan urgensi analisis sentimen untuk memahami persepsi pengguna secara lebih mendalam. Penelitian ini berfokus pada Identitas Kependudukan Digital (IKD), sebuah aplikasi identitas digital yang dikembangkan oleh pemerintah Indonesia. Tujuannya adalah untuk mengklasifikasikan ulasan pengguna terhadap aplikasi IKD ke dalam sentimen positif, netral, dan negatif menggunakan algoritma *random forest*. Dataset terdiri atas 28.134 ulasan pengguna dari Google Play Store, mencakup nama pengguna, teks ulasan, stempel waktu, dan penilaian bintang. Tahapan penelitian meliputi prapemrosesan data, pelabelan, penanganan nilai hilang, dan pemrosesan teks (pembersihan, tokenisasi, penghapusan *stopword*, dan *stemming*). Data dibagi menjadi dua bagian, yaitu 80% untuk data pelatihan dan 20% untuk data pengujian. Model terbaik menggunakan parameter: $max_depth = None$, $max_features = \log_2$, $min_samples_leaf = 1$, $min_samples_split = 2$, dan $n_estimators = 300$, dengan akurasi rata-rata sebesar 83,78%. Untuk mengatasi ketidakseimbangan kelas, diterapkan *synthetic minority oversampling technique* (SMOTE), yang menghasilkan peningkatan kinerja dengan akurasi mencapai 86,29%. Evaluasi sebelum SMOTE menunjukkan akurasi 83,85%, presisi 80,40%, *recall* 83,85%, dan nilai *F1-score* sebesar 81,73%. Setelah penerapan SMOTE, presisi meningkat menjadi 81,22%, sedangkan akurasi dan *recall* sedikit menurun menjadi masing-masing 80,86%, dengan *F1-score* sebesar 81,03%. Selain itu, dilakukan analisis tren sentimen menggunakan teknik *N-gram* (*unigram*, *bigram*, *trigram*) untuk mengidentifikasi topik dan permasalahan yang sering disebutkan oleh pengguna. Temuan ini mendukung tujuan penelitian dalam memberikan masukan perbaikan aplikasi yang selaras dengan kebutuhan pengguna serta meningkatkan pengalaman layanan digital secara keseluruhan.

KATA KUNCI — Analisis Sentimen, Klasifikasi Teks, Aplikasi IKD, *Random Forest*, SMOTE.

I. PENDAHULUAN

Identitas kependudukan memiliki peran penting dalam memastikan akses warga negara Indonesia terhadap berbagai layanan publik esensial, seperti layanan kesehatan, pendidikan, bantuan sosial, dan dokumen hukum. Sejalan dengan agenda transformasi digital, pemerintah Indonesia telah memperkenalkan Identitas Kependudukan Digital (IKD), yaitu identitas kependudukan berbasis digital yang dikembangkan dalam bentuk aplikasi seluler oleh Direktorat Jenderal Kependudukan dan Pencatatan Sipil (Ditjen Dukcapil). IKD merupakan versi digital dari Kartu Tanda Penduduk (KTP) yang memungkinkan pengguna untuk mengakses data pribadi melalui aplikasi di perangkat seluler. IKD bertujuan untuk meningkatkan efisiensi serta kemudahan akses terhadap berbagai layanan publik dan swasta tanpa memerlukan dokumen fisik, sekaligus mendorong inklusi digital dan ekonomi. Dengan menggunakan teknologi enkripsi tingkat lanjut, IKD menjamin keamanan data dan menawarkan berbagai manfaat, seperti kemudahan akses layanan, efisiensi administrasi, serta pemerataan akses bagi seluruh lapisan masyarakat [1]. Dalam studi ini, istilah “IKD” digunakan untuk merujuk pada platform identifikasi digital tersebut. Lebih lanjut, IKD juga bertujuan untuk meningkatkan efisiensi administrasi melalui pengurangan birokrasi, penyederhanaan proses layanan, dan peningkatan keamanan data. Namun demikian, implementasi IKD masih menghadapi berbagai tantangan, antara lain rendahnya literasi digital di kalangan masyarakat, ketimpangan infrastruktur, prosedur aktivasi yang kompleks, serta antarmuka aplikasi yang kurang intuitif [2], [3].

Mengingat berbagai tantangan yang dihadapi, pemahaman terhadap persepsi dan umpan balik pengguna terhadap aplikasi IKD menjadi sangat penting. Dalam konteks ini, pemelajaran mesin (*machine learning*), khususnya analisis sentimen, memainkan peran yang krusial. Analisis sentimen merupakan teknik komputasi yang digunakan untuk mengidentifikasi dan mengklasifikasikan opini yang diekspresikan dalam teks ke dalam kategori positif, negatif, atau netral. Metode ini umum digunakan untuk memahami persepsi publik terhadap suatu produk, layanan, atau isu tertentu, serta mendukung pengambilan keputusan strategis melalui wawasan yang diperoleh dari analisis sentimen yang terdeteksi [4]. Dengan didukung oleh algoritma seperti *random forest*, analisis sentimen mampu mengenali pola dalam ulasan pengguna dan mengklasifikasikannya berdasarkan jenis sentimen. *Random forest* sendiri merupakan metode *ensemble learning* yang dikenal karena ketangguhan, kinerja prediktif yang tinggi, serta kemampuannya dalam menangani *dataset* berukuran besar. Ketika dikombinasikan dengan teknik penambangan teks (*text mining*), metode ini menjadi alat yang sangat efektif untuk mengekstraksi wawasan dari data berbasis teks [5]-[7].

Penelitian sebelumnya telah menunjukkan efektivitas berbagai model pemelajaran mesin dalam analisis sentimen. Sebagai contoh, sebuah studi terhadap ulasan hotel di Purwokerto yang menggunakan algoritma *random forest* berhasil mencapai akurasi sebesar 87,23% meskipun tanpa penerapan proses *stemming* [8]. Penelitian lain yang menganalisis umpan balik pengguna terhadap aplikasi IKD dengan menggunakan algoritma *naïve Bayes* memperoleh

tingkat akurasi sebesar 85,06% [9]. Selain itu, studi komparatif yang dilakukan terhadap analisis sentimen pada data terkait metaverse menunjukkan bahwa algoritma *random forest* yang dikombinasikan dengan *synthetic minority oversampling technique* (SMOTE) memberikan hasil yang lebih baik dibandingkan dengan *support vector machine* (SVM), dengan tingkat akurasi mencapai 91% [10].

Meskipun hasil-hasil tersebut menunjukkan potensi yang menjanjikan, sebagian besar penelitian sebelumnya masih berfokus pada domain yang berbeda atau belum mempertimbangkan permasalahan ketidakseimbangan data (*data imbalance*). Oleh karena itu, penelitian ini diperlukan untuk menjembatani kesenjangan tersebut dengan menawarkan kombinasi teknik yang dirancang secara spesifik untuk konteks layanan publik di Indonesia. Penelitian ini memberikan kontribusi baru dengan memanfaatkan *dataset* berskala besar dari dunia nyata (sebanyak 28.134 ulasan), yang dikumpulkan langsung dari halaman aplikasi IKD di Play Store, menerapkan kombinasi algoritma *random forest* dan SMOTE untuk menangani ketidakseimbangan kelas, serta melakukan analisis pola sentimen secara mendalam menggunakan model *N-gram* (*unigram*, *bigram*, *trigram*). Selain itu, fokus utama pada ulasan pengguna berbahasa Indonesia meningkatkan relevansi lokal dari model yang dikembangkan. Studi ini tidak hanya mengevaluasi kinerja *random forest* dalam klasifikasi sentimen, tetapi juga mengungkap isu-isu yang paling sering dibahas serta fitur-fitur yang paling dihargai dalam aplikasi IKD berdasarkan umpan balik pengguna. Temuan yang diperoleh diharapkan dapat memberikan masukan berharga bagi perumusan kebijakan pemerintah, pengembangan aplikasi, serta mendukung penyediaan layanan publik digital yang lebih inklusif dan responsif di Indonesia.

II. PENELITIAN TERKAIT

Perkembangan analisis sentimen dalam konteks layanan publik digital makin mendapatkan perhatian di ranah akademik, terutama dalam mengevaluasi umpan balik pengguna terhadap aplikasi berbasis *mobile*. Seiring dengan meningkatnya penggunaan sistem identitas digital seperti IKD di Indonesia, pemahaman terhadap sentimen pengguna menjadi aspek penting untuk memastikan adopsi yang berhasil serta perbaikan berkelanjutan dari platform tersebut. Beberapa penelitian terdahulu telah mengkaji klasifikasi sentimen dengan menerapkan berbagai teknik pemelajaran mesin, strategi prapemrosesan tekstual, serta metrik evaluasi kinerja. Bagian ini mengulas studi-studi relevan yang menjadi landasan bagi penelitian ini, dengan menyoroti metode, temuan, serta keterbatasan dari penelitian sebelumnya yang menganalisis sentimen pada aplikasi identitas digital maupun layanan *e-government* sejenis.

Sebuah penelitian relevan pada tahun 2024 berfokus pada analisis sentimen terhadap ulasan pengguna aplikasi IKD di Google Play Store untuk memahami tanggapan publik dan memberikan masukan bagi pengembang [9]. Sebanyak 1.000 ulasan dikumpulkan melalui *web scraping* menggunakan Google Colab, kemudian diklasifikasikan secara otomatis berdasarkan pemeringkatan pengguna (*user rating*) sebagai sentimen positif atau negatif, dengan verifikasi manual untuk memastikan akurasi. Proses prapemrosesan mencakup pembersihan, *case folding*, tokenisasi, penghapusan *stopword*, dan penyaringan panjang token, dilanjutkan dengan pembobotan fitur menggunakan *term frequency-inverse document frequency* (TF-IDF). Untuk mengatasi

ketidakseimbangan kelas, digunakan metode SMOTE. Klasifikasi dilakukan dengan algoritma *naive Bayes*, menggunakan 90% data pelatihan dan 10% data pengujian, dan dievaluasi dengan *confusion matrix*. Hasil menunjukkan akurasi 85,06%, presisi 80,31%, dan *recall* 92,89%. Sebanyak 619 dari 1.000 ulasan dikategorikan sebagai sentimen negatif, dengan kata kunci dominan seperti “sulit”, “tolong”, “dokumen”, “buka”, dan “digital”, yang mencerminkan ketidakpuasan terhadap kemudahan dan fungsionalitas aplikasi. Kekuatan penelitian ini terletak pada penerapan SMOTE dan prapemrosesan yang komprehensif. Namun, terdapat beberapa keterbatasan, mencakup tidak digunakannya *stemming*, hanya satu algoritma klasifikasi yang diterapkan, serta ukuran *dataset* yang kecil. Penelitian lanjutan disarankan untuk mengeksplorasi algoritma lain dan memperluas *dataset* guna meningkatkan generalisasi hasil [9].

Studi relevan lainnya, pada tahun 2023, menganalisis sentimen publik terhadap aplikasi IKD, inisiatif pemerintah untuk menggantikan e-KTP dengan versi digital [11]. Data diperoleh dari komentar pengguna di Facebook selama 16 Februari hingga 10 Maret 2023 melalui teknik *comment crawling* dengan kata kunci “Identitas Kependudukan Digital”. Prapemrosesan mencakup pembersihan, *case folding*, tokenisasi, normalisasi, *filtering*, dan *stemming* menggunakan Python. Fitur teks diberi bobot menggunakan TF-IDF dan klasifikasi sentimen dilakukan menggunakan algoritma SVM dalam pendekatan multikelas (positif, negatif, netral). Evaluasi model dilakukan menggunakan *confusion matrix* dengan metrik akurasi, presisi, *recall*, dan *F1-score*, serta visualisasi dengan *word cloud*. Dari 902 komentar, 78,27% bersentimen negatif, 12,97% netral, dan hanya 8,76% positif. Model SVM mencatatkan akurasi 77%. Istilah yang sering muncul dalam komentar negatif, seperti “data” dan “proyek”, menunjukkan kekhawatiran publik terhadap kebocoran data dan skeptisisme terhadap program pemerintah. Kekuatan penelitian ini terletak pada penggunaan SVM yang efektif dalam analisis sentimen serta data Facebook yang mencerminkan opini waktu nyata. Namun, studi ini memiliki keterbatasan, yaitu cakupan data yang hanya bersumber dari satu platform serta adanya potensi bias data. Sementara itu, akurasi 77% menunjukkan perlunya peningkatan model dan seleksi fitur [11].

Studi lanjutan pada tahun 2024 menganalisis sentimen pengguna terhadap aplikasi IKD di Google Play Store menggunakan algoritma *k-nearest neighbor* (KNN) [12]. Tujuannya adalah mengklasifikasikan sentimen menjadi positif, negatif, atau netral serta mengevaluasi persepsi publik terhadap transformasi identitas digital di Indonesia. Data dikumpulkan melalui *web scraping*, lalu diproses dengan teknik *natural language processing* (NLP), seperti *case folding*, tokenisasi, pembersihan, penghapusan *stopword*, *stemming*, dan pelabelan, berdasarkan skor ulasan. Fitur teks diberi bobot menggunakan TF-IDF untuk memperkuat representasi dokumen. Klasifikasi dilakukan dengan KNN, diuji pada berbagai nilai k dan rasio pelatihan-pengujian. Konfigurasi terbaik dicapai pada $k = 13$ dan pembagian 80:20, dengan akurasi 83%, presisi 82%, *recall* 83%, *F1-score* 81%, dan *area under curve* (AUC) mikro 0,92. Model ini efektif untuk sentimen positif dan negatif, tetapi lemah dalam klasifikasi netral, ditandai dengan presisi 0,50, *recall* 0,01, dan *F1-score* 0,01, akibat jumlah sampel netral yang rendah. Sebanyak 66,75% ulasan tergolong negatif, mencerminkan ketidakpuasan pengguna terkait *usability* dan fungsionalitas aplikasi. Kekuatan penelitian ini terletak pada penggunaan KNN yang efektif, pemanfaatan TF-IDF, serta

evaluasi mendalam terhadap berbagai skenario pelatihan. Rekomendasi yang dihasilkan dapat membantu pengembang meningkatkan aplikasi. Namun, studi ini masih memiliki keterbatasan, yang mencakup ketidakseimbangan kelas, beban komputasi tinggi KNN untuk *dataset* besar, serta belum digunakannya analisis hubungan semantik dalam teks [12].

Tinjauan terhadap studi sebelumnya menunjukkan beberapa keterbatasan metodologis dan analitis. Salah satu studi menggunakan *naïve Bayes* dengan SMOTE untuk mengatasi ketidakseimbangan data, tetapi hanya mengandalkan satu algoritma tanpa perbandingan kinerja. Studi lain menggunakan SVM untuk klasifikasi tiga kelas sentimen, tetapi terbatas pada data dari Facebook, yang berpotensi menimbulkan bias lintas platform. Sementara itu, penelitian dengan algoritma KNN mengevaluasi berbagai nilai k , tetapi tetap kesulitan dalam mengklasifikasikan sentimen netral akibat distribusi data yang tidak seimbang.

Penelitian ini bertujuan mengatasi keterbatasan tersebut melalui pendekatan yang lebih komprehensif, dengan menerapkan algoritma *random forest* yang dikenal andal dalam pengolahan data teks dan tahan terhadap *overfitting*. Model dievaluasi sebelum dan sesudah penerapan SMOTE untuk mengukur dampaknya terhadap kinerja klasifikasi. Data diambil dari ulasan pengguna di Google Play Store guna memastikan cakupan sentimen yang lebih luas dan representatif. Prapemrosesan dilakukan secara menyeluruh, termasuk teknik normalisasi dan analisis *N-gram*, sehingga mampu memperkaya pemahaman pola sentimen publik terhadap aplikasi IKD.

III. METODOLOGI

Penelitian ini mengikuti metodologi terstruktur yang terdiri atas beberapa tahapan utama dalam menganalisis sentimen pengguna terhadap aplikasi IKD. Tahap pertama adalah pengumpulan data, yaitu pengumpulan ulasan dari halaman aplikasi IKD di Google Play Store menggunakan teknik *crawling* yang diimplementasikan dengan Python dan Google Play Store Scraper. Data yang dikumpulkan mencakup *username*, *review*, *date*, dan *score*, yang merepresentasikan identitas pengguna, isi ulasan, tanggal publikasi, dan penilaian aplikasi (skala 1–5).

Tahap berikutnya adalah prapemrosesan data, yang bertujuan memastikan kualitas dan kesiapan data sebelum dimodelkan dengan pembelajaran mesin. Tiga pendekatan utama dalam pembelajaran mesin adalah pembelajaran terbimbing (*supervised learning*), pembelajaran takterbimbing (*unsupervised learning*), dan pembelajaran kukuh (*reinforcement learning*)—masing-masing digunakan untuk klasifikasi, pengelompokan, dan pengambilan keputusan berbasis umpan balik [13]. Langkah awal prapemrosesan adalah pelabelan data, yang dilakukan secara manual oleh ahli untuk memberi label *positive*, *negative*, atau *neutral*. Ulasan *neutral* yang bersifat spam atau tidak relevan dikeluarkan dari *dataset*. Selanjutnya, dilakukan pembersihan dan normalisasi teks, meliputi pembersihan, tokenisasi, penghapusan *stopword*, *stemming*, dan penanganan nilai hilang untuk menghilangkan derau dan menstandarkan data. Terakhir, *dataset* dibagi menjadi 80% data pelatihan dan 20% data pengujian, guna melatih model dan mengevaluasi kemampuannya dalam menggeneralisasi pola sentimen.

Untuk mengubah data teks hasil prapemrosesan menjadi format numerik yang sesuai untuk pemodelan pembelajaran mesin, digunakan teknik TF-IDF. Metode ini memberi bobot

pada kata berdasarkan frekuensinya dalam dokumen dan kelangkaannya di seluruh korpus, sehingga membantu menonjolkan istilah penting dan mengurangi pengaruh kata umum. Representasi ini esensial dalam meningkatkan akurasi klasifikasi sentimen [14]. Pemodelan dilakukan menggunakan algoritma *random forest*, teknik *ensemble learning* yang menggabungkan banyak pohon keputusan dan menentukan prediksi akhir melalui pilihan terbanyak. Algoritma ini dikenal tahan terhadap *overfitting* dan memiliki akurasi tinggi, menjadikannya lebih andal dibanding SVM atau *naïve Bayes* dalam menangani *dataset* kompleks [15]. Proses pemodelan disempurnakan dengan menerapkan teknik *grid search* guna menemukan kombinasi *hyperparameter* terbaik untuk meningkatkan kinerja model. Pemodelan dilakukan pada dua skenario, yakni menggunakan data asli dan data yang telah ditangani dengan metode SMOTE. SMOTE diterapkan untuk menyeimbangkan distribusi kelas, terutama ketika terdapat dominasi satu kelas sentimen yang signifikan dibandingkan kelas lainnya, yang dapat menyebabkan penurunan kinerja model. SMOTE adalah metode *oversampling* yang dirancang untuk mengatasi masalah ketidakseimbangan kelas dalam data pelatihan yang digunakan oleh algoritma klasifikasi [16].

Evaluasi dilakukan menggunakan data pengujian untuk membandingkan kinerja model, termasuk akurasi, presisi, dan *recall*, sebelum dan sesudah penerapan SMOTE [17]. Hasilnya membantu mengidentifikasi ekspresi umum serta isu yang sering muncul dalam umpan balik pengguna.

A. PENGUMPULAN DATA

Data yang digunakan dalam penelitian ini diperoleh dari platform Google Play Store, yang berfungsi sebagai sumber utama ulasan pengguna terkait aplikasi IKD. Google Play Store tidak hanya menyediakan informasi mengenai aplikasi, tetapi juga memuat ulasan yang mencerminkan pengalaman dan persepsi pengguna, sehingga menjadikannya sumber data yang relevan untuk analisis sentimen. Aplikasi IKD tersedia di Play Store sebagai aplikasi berbasis Android dengan versi 1.2.2, yang terakhir diperbarui pada tanggal 24 Mei 2023. Aplikasi ini pertama kali dirilis pada 3 Juni 2022 oleh Direktorat Jenderal Kependudukan dan Pencatatan Sipil (Ditjen Dukcapil) Kementerian Dalam Negeri (Kemendagri). Aplikasi tersebut memiliki ukuran *file* sebesar 22 MB dan kompatibel dengan perangkat Android versi 5.0 ke atas. Hingga saat ini, aplikasi IKD telah diunduh lebih dari 10 juta kali dan diklasifikasikan sebagai aplikasi yang cocok digunakan oleh pengguna berusia 3 tahun ke atas [18].

Penelitian ini mengumpulkan sebanyak 28.134 ulasan melalui proses *crawling*. Data yang diekstraksi mencakup beberapa elemen penting, yaitu nama pengguna, teks ulasan, tanggal ulasan, dan skor penilaian yang diberikan pengguna dalam rentang 1 hingga 5. Distribusi skor ulasan tersebut menunjukkan adanya polarisasi signifikan dalam opini pengguna, yang mencerminkan keberagaman pengalaman serta persepsi terhadap aplikasi IKD.

Ulasan dengan skor 1, yang mencerminkan tingkat ketidakpuasan tertinggi, berjumlah 12.648 ulasan atau sekitar 44,94% dari keseluruhan data. Hal ini menunjukkan bahwa banyak pengguna merasa tidak puas terhadap penggunaan aplikasi IKD. Sebaliknya, ulasan dengan skor 5, yang merepresentasikan tingkat kepuasan tertinggi, tercatat sebanyak 11.421 ulasan atau sekitar 40,60% dari keseluruhan ulasan. Angka-angka tersebut mengindikasikan bahwa meskipun terdapat jumlah ulasan negatif yang signifikan,

aplikasi IKD juga mendapatkan apresiasi dari sejumlah besar pengguna yang merasa puas.

Ulasan dengan skor 2, 3, dan 4 memiliki distribusi yang lebih kecil. Secara rinci, skor 2 mencakup 1.762 ulasan (6,26%), skor 3 sebanyak 1.352 ulasan (4,81%), dan skor 4 sejumlah 951 ulasan (3,38%). Dominasi skor ekstrem, yaitu 1 dan 5, menunjukkan bahwa persepsi pengguna terhadap aplikasi IKD cenderung bersifat sangat positif atau sangat negatif, dengan jumlah ulasan yang relatif sedikit pada kisaran moderat. Data ulasan yang telah dikumpulkan ini menjadi landasan penting dalam analisis sentimen karena memberikan gambaran mendalam mengenai pengalaman pengguna dan persepsi publik terhadap aplikasi IKD.

B. PRAPEMROSESAN DATA

Prapemrosesan data merupakan tahapan krusial dalam pemrosesan data yang bertujuan untuk meningkatkan kualitas data sebelum digunakan dalam analisis atau pelatihan model. Tahapan ini berfungsi untuk mengubah data mentah menjadi format yang lebih terstruktur dan bersih, sehingga memungkinkan proses analisis dilakukan secara efisien serta meningkatkan kinerja model. Prapemrosesan berperan penting dalam menghilangkan informasi yang tidak relevan, mengatasi inkonsistensi atau kekurangan data, serta menyiapkan fitur-fitur yang relevan untuk analisis selanjutnya. Dalam penelitian ini, tahapan prapemrosesan data mencakup langkah-langkah sebagai berikut.

1) PELABELAN

Pada tahap ini, setiap ulasan pengguna diberi label sentimen berdasarkan skor *rating* yang diberikan. *Rating* tinggi (4–5) diklasifikasikan sebagai sentimen positif, sedangkan *rating* rendah (1–2) dikategorikan sebagai sentimen negatif. *Rating* dengan skor 3 diklasifikasikan sebagai sentimen netral. Proses ini bertujuan untuk menyediakan variabel target yang akan digunakan dalam klasifikasi sentimen.

Dari keseluruhan data yang dianalisis, terdapat 11.522 ulasan dengan sentimen negatif, 1.113 ulasan dengan sentimen netral, dan 11.251 ulasan dengan sentimen positif. Distribusi ini menunjukkan bahwa jumlah ulasan negatif dan positif relatif seimbang, sedangkan jumlah ulasan netral jauh lebih sedikit. Hal ini mengindikasikan bahwa pengguna cenderung menyampaikan pendapat yang kuat, baik dalam bentuk apresiasi maupun keluhan, dibandingkan memberikan umpan balik yang bersifat netral atau tidak memihak.

2) PENANGANAN NILAI HILANG (MISSING VALUES HANDLING)

Penanganan nilai hilang dilakukan untuk mendeteksi entri yang tidak valid atau kosong pada kolom teks ulasan. Fungsi seperti `isnull()` dan metode `str.strip()` digunakan untuk mengidentifikasi nilai *NaN* atau *string* kosong. Entri yang tidak valid, seperti teks kosong atau *string* yang hanya terdiri atas spasi, dihapus menggunakan metode `dropna()` dan filter tambahan. Setelah proses pembersihan selesai, jumlah nilai yang hilang dihitung kembali untuk memastikan tidak ada entri kosong yang tersisa. Proses ini menjamin bahwa *dataset* yang digunakan sepenuhnya valid, dengan hasil akhir menunjukkan nol nilai yang hilang.

3) PRAPEMROSESAN TEKS

Proses ini dilakukan untuk menyaring dan mempersiapkan data teks sebelum analisis lebih lanjut. Tahapan ini mencakup beberapa langkah penting. Pertama, dilakukan proses pembersihan untuk menghapus karakter khusus, angka, tanda

baca, serta elemen lain yang tidak relevan. Selanjutnya, dilakukan tokenisasi untuk memecah teks ulasan menjadi unit kata individu guna memfasilitasi analisis yang lebih terstruktur. Setelah itu, diterapkan penghapusan *stopword* untuk menghilangkan kata-kata umum yang tidak memberikan makna signifikan, seperti “yang”, “dan”, atau “adalah”. Langkah terakhir adalah *stemming*, yang menggunakan pustaka Sastrawi untuk mengubah kata-kata menjadi bentuk dasarnya; misalnya, kata “berlari” akan dikonversi menjadi “lari”. Seluruh tahapan ini menghasilkan data teks yang bersih dan terstruktur, sehingga siap untuk tahap ekstraksi fitur dan klasifikasi selanjutnya.

4) PEMBAGIAN DATA

Dataset dibagi menjadi dua bagian, yaitu 80% untuk pelatihan dan 20% untuk pengujian). Dari keseluruhan 28.134 data ulasan yang tersedia, sekitar 22.507 data digunakan untuk proses pelatihan model (*training set*), sedangkan 5.627 data sisanya digunakan untuk menguji kinerja model (*test set*).

Pembagian ini bertujuan untuk memastikan bahwa model memiliki cukup data untuk mempelajari pola-pola dari ulasan pengguna, sekaligus menyediakan data yang terpisah untuk mengevaluasi kemampuan generalisasi model dalam mengklasifikasikan sentimen. Strategi pembagian data ini penting untuk menghindari *overfitting* dan memberikan gambaran yang lebih akurat terhadap kinerja model dalam konteks data yang belum pernah dilihat sebelumnya.

5) TF-IDF

Proses TF-IDF mengubah data teks menjadi representasi numerik dalam bentuk *sparse matrix*, yaitu X_{train_tfidf} untuk data pelatihan dan X_{test_tfidf} untuk data pengujian. Matriks tersebut mencakup 10.000 fitur yang merepresentasikan kata-kata dalam *dataset* dengan bobot berdasarkan tingkat kepentingannya. Matriks TF-IDF yang dihasilkan ini disiapkan sebagai masukan untuk model pembelajaran mesin pada tahap analisis dan klasifikasi sentimen. Dengan menggunakan representasi ini, model dapat memproses data teks dalam format numerik yang dapat dihitung, sehingga memungkinkan pelatihan algoritma klasifikasi untuk mengenali pola-pola sentimen secara lebih efektif dan akurat.

C. PEMODELAN RANDOM FOREST

Tahap pemodelan *random forest* bertujuan untuk membangun model klasifikasi yang optimal dalam menganalisis sentimen pengguna terhadap aplikasi IKD. Proses pemodelan ini dilakukan melalui dua pendekatan utama: tanpa penerapan SMOTE dan dengan penerapan SMOTE pada data pelatihan. *Random forest* merupakan metode klasifikasi yang membangun sejumlah pohon keputusan berdasarkan sampel data yang diambil secara acak. Proses ini melibatkan tiga tahapan utama, di antaranya *bootstrap sampling*, yaitu pengambilan sampel data secara acak dengan pengembalian (*sampling with replacement*); *random sub-setting*, yaitu pemilihan fitur secara acak saat membangun setiap pohon keputusan; dan *aggregating*, yaitu penggabungan prediksi dari seluruh pohon untuk menentukan keluaran akhir. Metode ini dikenal efektif dalam menangani data nonlinear serta dalam mengidentifikasi variabel-variabel paling relevan pada analisis data yang kompleks [19].

Dalam pendekatan pemodelan tanpa menggunakan SMOTE, model dibangun dengan menerapkan metode *grid search* yang dikombinasikan dengan strategi *5-fold cross-validation* untuk menentukan konfigurasi *hyperparameter* yang optimal. Parameter yang disesuaikan mencakup jumlah

pohon ($n_{estimators}$), kedalaman maksimum dari pohon keputusan (max_depth), jumlah minimum sampel yang dibutuhkan untuk melakukan pemisahan ($min_samples_split$), jumlah minimum sampel pada setiap daun ($min_samples_leaf$), serta jumlah maksimum fitur yang digunakan ($max_features$). Metode *grid search* mengevaluasi 288 kombinasi parameter, yang menghasilkan 1.440 evaluasi model melalui proses *5-fold cross-validation*. Hasil akhir dari proses ini adalah serangkaian parameter yang telah dioptimalkan dan mampu memberikan rata-rata akurasi tertinggi pada data pelatihan.

Dalam strategi pemodelan, SMOTE diterapkan untuk mengatasi ketidakseimbangan kelas pada data pelatihan dengan menghasilkan sampel sintetis pada kelas minoritas. Teknik ini membantu menyeimbangkan distribusi kelas, sehingga model dapat belajar lebih efektif dari semua kategori data. Setelah SMOTE diterapkan, data yang telah dicuplik ulang (*resampled*) digunakan untuk membangun model *random forest* dengan pendekatan *grid search* dan *cross-validation*. Tujuannya adalah meningkatkan akurasi klasifikasi, terutama pada kelas minoritas, dan menghasilkan model yang lebih seimbang dan andal dibandingkan model tanpa SMOTE.

D. PENGUJIAN RANDOM FOREST

Tahap evaluasi *random forest* bertujuan menilai efektivitas model dalam mengklasifikasikan data pengujian, baik sebelum maupun sesudah penerapan SMOTE. Evaluasi dilakukan menggunakan metrik utama: akurasi, presisi, *recall*, dan *F1-score*. Akurasi mengukur ketepatan keseluruhan, presisi menunjukkan keandalan prediksi positif, *recall* menilai sensitivitas model terhadap sentimen positif, dan *F1-score* menyeimbangkan presisi dan *recall*. Penghitungan metrik ini didasarkan pada *confusion matrix*, yang memetakan hasil klasifikasi ke dalam empat kategori: *true positive* (TP), *true negative* (TN), *false positive* (FP), dan *false negative* (FN) [20]. Evaluasi ini membandingkan dua skenario untuk menilai kontribusi SMOTE dalam meningkatkan kinerja model terhadap data yang tidak seimbang serta memastikan efektivitasnya dalam analisis sentimen ulasan aplikasi IKD.

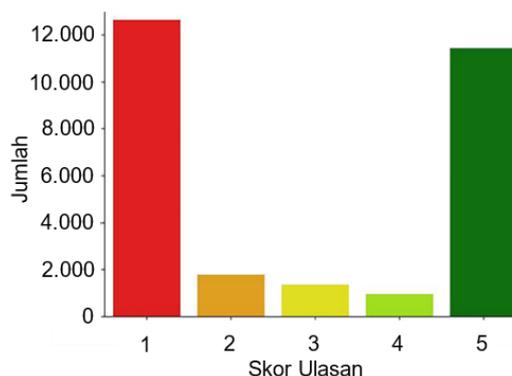
IV. HASIL

A. PENGUMPULAN DATA

Proses pengumpulan data dilakukan menggunakan metode *scraping* dengan memanfaatkan bahasa pemrograman Python serta *google-play-scraper*. Sebanyak 28.134 ulasan berhasil dikumpulkan, mencakup beragam pendapat pengguna terkait pengalaman dalam menggunakan aplikasi. Setiap ulasan memuat informasi seperti nama pengguna, teks ulasan, tanggal ulasan, serta skor (penilaian bintang) yang diberikan oleh pengguna, dengan rentang nilai antara 1 hingga 5.

Berdasarkan distribusi skor (Gambar 1), data ulasan didominasi oleh dua kelompok utama, yaitu skor 1 dan skor 5. Ulasan dengan skor 1, yang mencerminkan tingkat ketidakpuasan tertinggi, berjumlah 12.648 entri atau sekitar 44,94% dari total *dataset*. Hal ini menunjukkan bahwa sebagian besar pengguna mengalami pengalaman yang kurang memuaskan saat menggunakan aplikasi IKD. Di sisi lain, ulasan dengan skor 5, yang merepresentasikan tingkat kepuasan tertinggi, juga mencakup jumlah yang signifikan, yaitu sebanyak 11.421 ulasan atau sekitar 40,60% dari keseluruhan data. Hal ini mengindikasikan bahwa meskipun banyak pengguna merasa tidak puas, terdapat pula sejumlah besar pengguna yang sangat mengapresiasi aplikasi tersebut.

Kelompok skor lainnya menunjukkan jumlah yang jauh lebih kecil. Ulasan dengan skor 2 berjumlah 1.762 entri (sekitar



Gambar 1. Distribusi *dataset* berdasarkan skor.

6,26%), skor 3 sebanyak 1.352 ulasan (sekitar 4,81%), dan skor 4 terdiri atas 951 ulasan (sekitar 3,38%). Distribusi ini menunjukkan bahwa opini pengguna terhadap aplikasi sangat terpolarisasi—sebagian besar pengguna memberikan ulasan yang sangat positif atau sangat negatif—sedangkan ulasan dengan penilaian sedang (skor 2, 3, dan 4) relatif jarang ditemukan.

B. PRAPEMROSESAN DATA

Pada tahap prapemrosesan, setiap ulasan pengguna diberi label sentimen—positif, negatif, atau netral—berdasarkan skor *rating* yang diberikan. *Rating* tinggi (4–5) diasumsikan mencerminkan sentimen positif, sementara *rating* rendah (1–2) dikategorikan sebagai sentimen negatif. Sementara itu, *rating* dengan skor 3 diklasifikasikan sebagai sentimen netral. Tujuan dari proses pelabelan ini adalah untuk mempersiapkan data sebagai variabel target dalam proses klasifikasi sentimen.

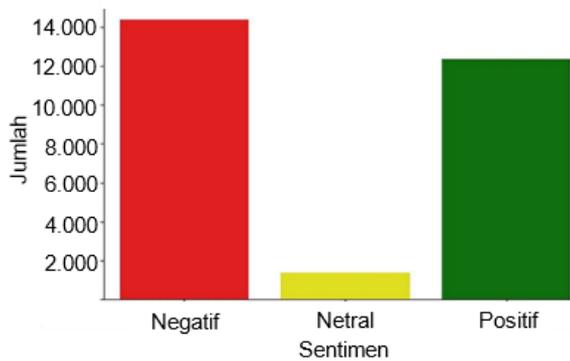
Berdasarkan hasil pelabelan yang ditunjukkan pada Gambar 2, sentimen negatif—dilabeli sebagai -1—mencakup 14.410 ulasan, yang menunjukkan bahwa sekitar 51,2% dari data yang dianalisis mencerminkan ketidakpuasan pengguna terhadap aplikasi IKD. Sebaliknya, sentimen positif—dilabeli sebagai 1—terdapat pada 12.372 ulasan, atau sekitar 44% dari keseluruhan data, yang menunjukkan bahwa sejumlah besar pengguna memiliki pengalaman yang baik dengan aplikasi tersebut. Ulasan netral—dilabeli sebagai 0—berjumlah hanya 1.352 entri, atau sekitar 4,8%. Ulasan netral ini umumnya tidak mengandung muatan emosional yang kuat dan sering kali berisi komentar yang tidak relevan, tautan, atau spam, sehingga kurang berguna untuk analisis lanjutan.

Distribusi sentimen ini mengindikasikan adanya kecenderungan opini pengguna yang terpolarisasi, yaitu mayoritas menyampaikan kepuasan atau ketidakpuasan secara tegas dan hanya sedikit yang bersifat netral. Proses pelabelan ini memungkinkan analisis yang lebih mendalam terhadap faktor-faktor yang memengaruhi sentimen pengguna, yang dapat menjadi dasar penting dalam memahami persepsi publik terhadap aplikasi serta memberikan wawasan yang dapat ditindaklanjuti untuk pengembangan di masa mendatang.

C. PEMODELAN RANDOM FOREST

1) PARAMETER YANG DIGUNAKAN

Pada tahap ini, model *random forest* dibangun menggunakan teknik *grid search* dengan metode *5-fold cross-validation*. Berdasarkan hasil eksperimen yang dilakukan, model dengan kinerja terbaik diperoleh melalui kombinasi parameter sebagai berikut: kedalaman pohon tidak dibatasi ($max_depth = None$), jumlah fitur maksimum yang digunakan



Gambar 2. Hasil pelabelan sentimen.

pada setiap pemisahan *node* adalah \log_2 , jumlah minimum sampel pada daun adalah 1 ($min_samples_leaf = 1$), jumlah minimum sampel untuk memisahkan *node* adalah 2 ($min_samples_split = 2$), dan jumlah keseluruhan pohon yang digunakan sebanyak 300 ($n_estimators = 300$). Model optimal ini berhasil mencapai rata-rata akurasi sebesar 83,78%. Berdasarkan hasil tersebut, kombinasi parameter $\{ 'max_depth': None, 'max_features': 'log2', 'min_samples_leaf': 1, 'min_samples_split': 2, 'n_estimators': 300 \}$ dipilih sebagai konfigurasi terbaik karena menghasilkan skor akurasi rata-rata tertinggi.

2) IMPLEMENTASI SMOTE

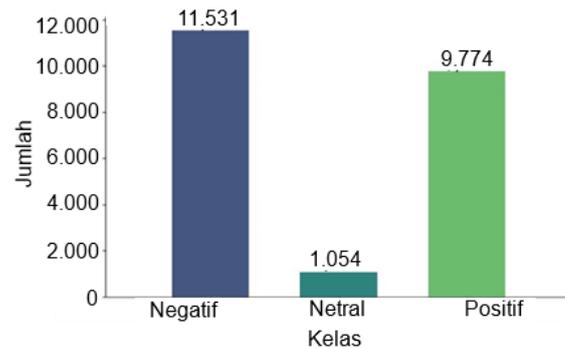
Dalam tugas klasifikasi, ketidakseimbangan kelas pada data pelatihan merupakan permasalahan umum, dengan jumlah sampel pada salah satu kelas jauh lebih sedikit dibandingkan dengan kelas lainnya. Ketidakseimbangan ini dapat berdampak negatif terhadap kinerja model, terutama dalam hal akurasi dan kemampuan generalisasi. Untuk mengatasi permasalahan tersebut, digunakan metode SMOTE yang berfungsi untuk meningkatkan jumlah sampel pada kelas minoritas dengan cara menghasilkan data sintesis. Dengan adanya penambahan data buatan ini, distribusi kelas menjadi lebih seimbang, sehingga model dapat belajar secara lebih efektif terhadap seluruh kelas yang ada.

Gambar 3 menunjukkan distribusi kelas sebelum penerapan SMOTE. Terlihat bahwa *dataset* memiliki ketidakseimbangan kelas yang cukup signifikan, dengan 11.531 sampel pada kelas *negative*, 9.774 sampel pada kelas *positive*, dan hanya 1.054 sampel pada kelas *neutral*. Ketimpangan representasi ini dapat menyebabkan prediksi yang bias, karena model cenderung kesulitan dalam mengidentifikasi atau mengklasifikasikan data dari kelompok minoritas, khususnya kelas *neutral*.

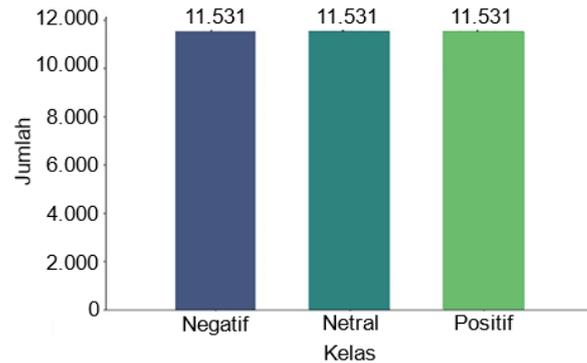
Untuk mengatasi permasalahan tersebut, dilakukan penerapan SMOTE terhadap data pelatihan, yaitu pada X_{train_tfidf} dan y_{train} . Proses ini menghasilkan dua variabel baru, yakni $X_{train_resampled}$, yang berisi fitur-fitur hasil pencuplikan ulang, dan $y_{train_resampled}$, yang berisi label yang sesuai. SMOTE bekerja dengan cara mengidentifikasi kelas-kelas yang kurang terwakili, kemudian menghasilkan sampel sintesis baru berdasarkan pola dari data yang ada.

Setelah penerapan SMOTE, dilakukan pemeriksaan ulang terhadap distribusi kelas pada $y_{train_resampled}$ guna memastikan bahwa setiap kelas memiliki jumlah sampel yang setara. Proses ini melibatkan penghitungan jumlah *instance* untuk setiap label kelas dalam $y_{train_resampled}$.

Hasilnya menunjukkan bahwa jumlah sampel untuk masing-masing kelas, yaitu -1 untuk *negative*, 0 untuk *neutral*, dan 1 untuk *positive*, menjadi seimbang, yaitu masing-masing



Gambar 3. Distribusi kelas sebelum penerapan SMOTE.



Gambar 4. Distribusi sentimen setelah penerapan SMOTE.

sebanyak 11.531 sampel (Gambar 4). Keseimbangan ini sangat penting untuk mencegah model menjadi bias terhadap kelas mayoritas, serta meningkatkan kemampuannya dalam mengenali pola dari seluruh kategori sentimen. Secara keseluruhan, langkah ini diharapkan dapat meningkatkan akurasi model dan kemampuannya untuk melakukan generalisasi terhadap data baru yang belum pernah dilihat sebelumnya.

D. PENGUJIAN RANDOM FOREST

Dalam penelitian ini, prosedur pengujian dilakukan dengan membandingkan kinerja model dalam dua skenario yang berbeda, yaitu tanpa penerapan SMOTE dan dengan penerapan SMOTE. Proses evaluasi mempertimbangkan metrik kinerja utama, seperti akurasi, presisi, *recall*, dan *F1-score*, guna memastikan konsistensi dan ketangguhan model dalam berbagai kondisi data. Tujuan dari perbandingan ini adalah untuk menganalisis dampak penerapan SMOTE terhadap kemampuan model dalam menangani ketidakseimbangan kelas secara efektif.

Tabel I menyajikan hasil evaluasi model sebelum dan sesudah penerapan SMOTE. Evaluasi model sebelum penerapan SMOTE menunjukkan nilai akurasi sebesar 0,838462, yang berarti model mampu mengklasifikasikan dengan benar sekitar 83,85% dari seluruh ulasan. Nilai presisi mencapai 0,803984 mengindikasikan bahwa sekitar 80,40% dari ulasan yang diprediksi sebagai sentimen positif memang benar-benar positif. Sementara itu, nilai *recall* sebesar 0,838462 menunjukkan bahwa model telah berhasil mengidentifikasi sekitar 83,85% dari seluruh ulasan positif. Sebelum penerapan SMOTE, *F1-score* tercatat sebesar 0,817261, yang mencerminkan keseimbangan antara nilai presisi dan *recall*, dengan kinerja keseluruhan model berada pada kisaran 81,73%.

Setelah penerapan SMOTE, terjadi perubahan pada nilai-nilai metrik evaluasi. Akurasi model mengalami sedikit

TABEL I
HASIL EVALUASI MODEL

Metrik	Sebelum SMOTE	Setelah SMOTE
Akurasi	0,838462	0,808587
Presisi	0,803984	0,812209
Recall	0,838462	0,808587
F1-score	0,817261	0,810345

penurunan menjadi 0,808587, yang menunjukkan adanya sedikit penurunan kemampuan model dalam mengklasifikasikan ulasan secara tepat. Namun demikian, nilai presisi meningkat menjadi 0,812209, yang mengindikasikan bahwa model menjadi lebih akurat dalam mengklasifikasikan ulasan positif. Sebaliknya, *recall* mengalami penurunan menjadi 0,808587, menunjukkan bahwa model sedikit lebih sedikit dalam mengidentifikasi ulasan positif dibandingkan sebelum penerapan SMOTE. Terakhir, *F1-score* setelah penerapan SMOTE adalah sebesar 0,810345, yang meskipun menunjukkan sedikit penurunan dibandingkan nilai sebelumnya, tetap mencerminkan keseimbangan yang baik antara presisi dan *recall*.

E. HASIL ANALISIS

Bagian ini menyajikan hasil analisis sentimen yang dilakukan terhadap ulasan pengguna aplikasi IKD di Play Store. Analisis ini mencakup kinerja model *random forest*, pola-pola sentimen yang terdeteksi, serta kajian mendalam terhadap keluhan dan tingkat kepuasan pengguna.

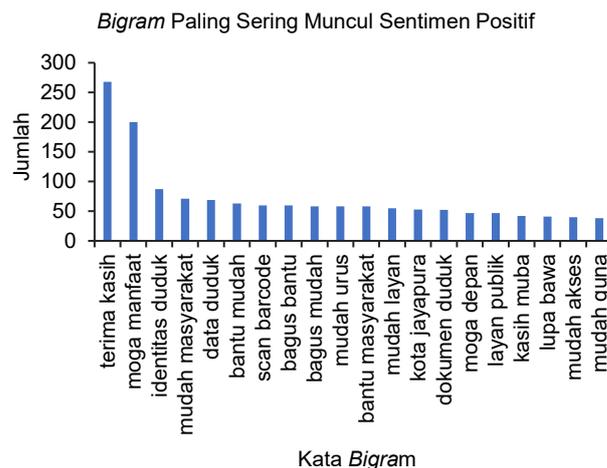
Hasil klasifikasi sentimen dibagi ke dalam tiga kategori, yaitu ulasan positif, negatif, dan netral. Untuk setiap kategori, diamati pola-pola yang muncul dalam umpan balik pengguna untuk mengidentifikasi tema-tema umum yang sering muncul.

Selain itu, dilakukan analisis kata untuk masing-masing kategori sentimen menggunakan metode *N-gram* dengan nilai $n = 1, 2, \text{ dan } 3$. *N-gram* merupakan teknik dalam NLP yang digunakan untuk mengekstraksi urutan kata yang sering muncul dalam teks.

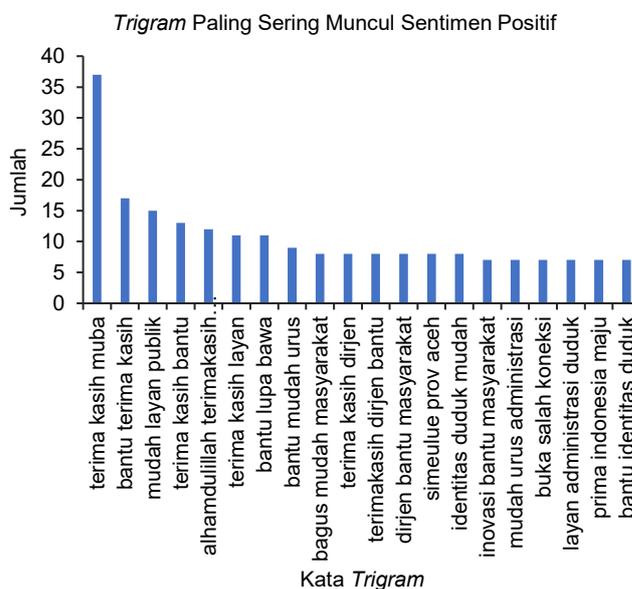
1. *Unigram* ($n = 1$) menganalisis kata per kata secara individual, sehingga memungkinkan identifikasi istilah tunggal yang paling sering digunakan dalam ulasan positif maupun negatif.
2. *Bigram* ($n = 2$) mengevaluasi pasangan kata yang sering muncul berurutan, membantu mengungkap frasa umum yang menggambarkan keluhan pengguna atau fitur yang paling diapresiasi.
3. *Trigram* ($n = 3$) memusatkan perhatian pada tiga kata yang muncul secara berurutan, memberikan konteks yang lebih kaya mengenai cara pengguna menyampaikan kritik atau pujian.

Dengan menganalisis kata dan frasa yang paling sering muncul dalam ulasan pengguna, dapat diperoleh wawasan yang lebih mendalam mengenai aspek-aspek aplikasi yang memerlukan perbaikan serta fitur-fitur yang dianggap memuaskan oleh pengguna. Informasi ini sangat penting bagi pengembang dalam melakukan peningkatan secara terarah berdasarkan umpan balik langsung dari pengguna. Hasil dari analisis *N-gram* ini akan disajikan secara rinci guna mendukung pemahaman yang lebih jelas terhadap persepsi pengguna terhadap aplikasi IKD.

Gambar 5 menggambarkan hasil implementasi *bigram* untuk kategori sentimen positif. Implementasi *bigram* pada sentimen positif mengungkap kombinasi kata seperti “terima kasih” (268 kemunculan), “moga manfaat” (200 kemunculan),



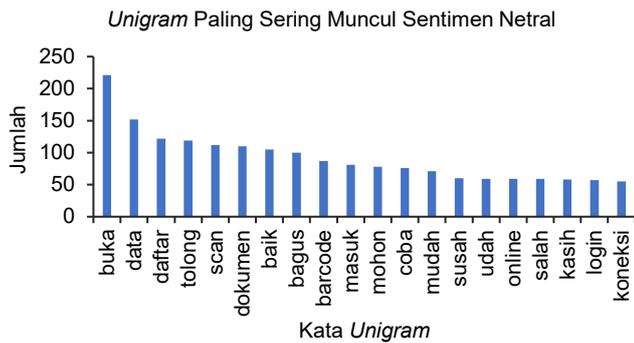
Gambar 5. Frekuensi kata berdasarkan *bigram* untuk sentimen positif.



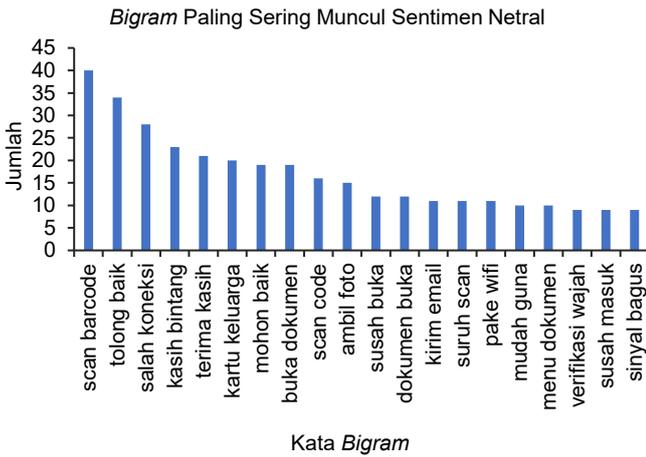
Gambar 6. Frekuensi kata berdasarkan *trigram* untuk sentimen positif.

dan “mudah masyarakat” (71 kemunculan), mencerminkan apresiasi dan kemudahan yang dirasakan oleh pengguna. Frasa seperti “identitas duduk” (87 kemunculan), “data duduk” (69 kemunculan), dan “dokumen duduk” (52 kemunculan) menunjukkan kemudahan dalam layanan administrasi. Fitur teknologi seperti pemindaian *barcode* juga mendapat apresiasi, sebagaimana terlihat dari frasa “*scan barcode*” (60 kemunculan) dan “mudah akses” (40 kemunculan). Selain itu, frasa seperti “bagus bantu”, “bagus mudah”, dan “mudah guna” menandakan persepsi positif terhadap kualitas layanan serta kemudahan penggunaan aplikasi.

Gambar 6 menunjukkan hasil implementasi *trigram* untuk kategori sentimen positif. Implementasi *trigram* pada sentimen positif menampilkan frasa-frasa yang memperkuat kepuasan pengguna, seperti “terima kasih muba” (37 kemunculan), “bantu terima kasih” (17 kemunculan), dan “mudah layanan publik” (15 kemunculan), yang mencerminkan rasa syukur atas layanan yang mempermudah. Frasa seperti “alhamdulillah terimakasih dirjen” (12 kemunculan) dan “terima kasih dirjen” (8 kemunculan) menunjukkan apresiasi terhadap dukungan dari pihak otoritas terkait. Selain itu, frasa seperti “bantu lupa bawa” (11 kemunculan) dan “mudah urus administrasi” (7 kemunculan) menyoroti aspek praktis serta kemudahan layanan yang dirasakan, sedangkan frasa “inovasi bantu masyarakat” (7



Gambar 7. Frekuensi kata berdasarkan unigram untuk sentimen netral.

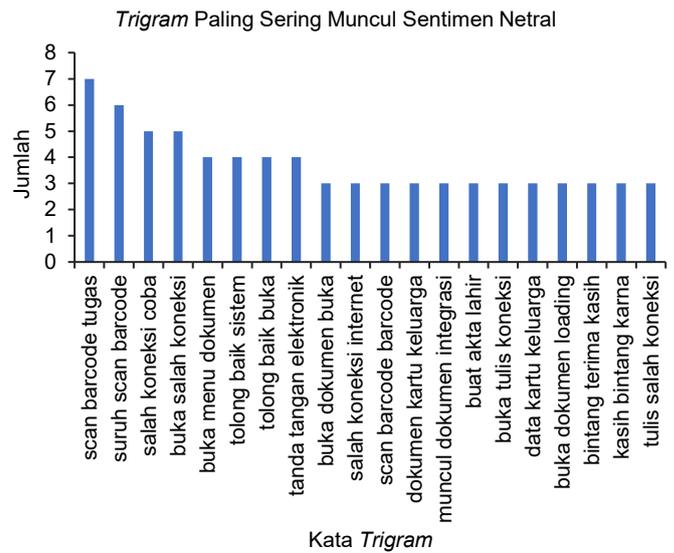


Gambar 8. Frekuensi kata berdasarkan bigram untuk sentimen netral.

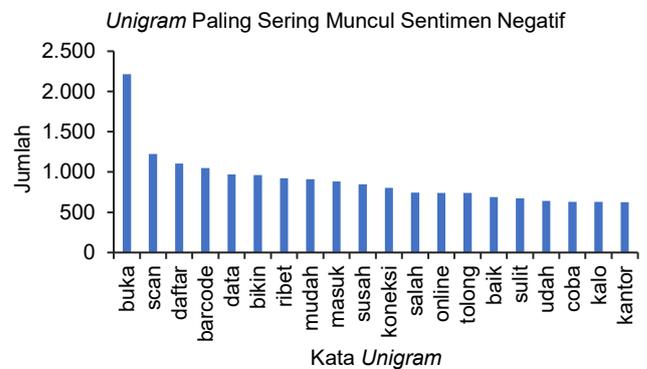
kemunculan) menggambarkan dampak positif dari inovasi terhadap masyarakat.

Gambar 7 menunjukkan hasil implementasi unigram untuk kategori sentimen netral. Implementasi unigram pada sentimen netral mengungkapkan kata-kata yang mencerminkan pengalaman pengguna yang beragam. Kata-kata seperti “buka” (221 kemunculan), “data” (152 kemunculan), dan “daftar” (122 kemunculan) menunjukkan fokus pada proses akses dan pendaftaran. Kata-kata seperti “tolong” (119 kemunculan) dan “mohon” (78 kemunculan) mengindikasikan adanya permintaan bantuan terkait kendala teknis. Istilah seperti “scan” (112 kemunculan), “barcode” (87 kemunculan), dan “login” (57 kemunculan) merujuk pada penggunaan teknologi dalam aplikasi, sedangkan kata-kata seperti “susah” (60 kemunculan), “salah” (59 kemunculan), dan “koneksi” (55 kemunculan) menyoroti tantangan teknis seperti kesalahan sistem atau permasalahan konektivitas.

Gambar 8 menunjukkan hasil implementasi bigram untuk kategori sentimen netral. Implementasi bigram pada sentimen netral mengungkapkan pengalaman pengguna yang beragam, mencerminkan perpaduan antara kepuasan dan kendala teknis. Frasa seperti “scan barcode” (40 kemunculan) menunjukkan penggunaan fitur yang cukup sering, meskipun berpotensi menimbulkan tantangan dalam penggunaannya. Frasa seperti “tolong baik” (34 kemunculan) dan “mohon baik” (19 kemunculan) mencerminkan permintaan pengguna terhadap layanan yang lebih lancar. Bigram seperti “salah koneksi” (28 kemunculan) dan “susah buka” (12 kemunculan) menunjukkan adanya permasalahan teknis, khususnya dalam hal akses dan proses login. Frasa seperti “kartu keluarga” (20 kemunculan) dan “verifikasi wajah” (9 kemunculan) menyoroti proses verifikasi dokumen yang sering digunakan, sedangkan frasa “pake wifi” (11 kemunculan) dan “sinyal bagus” (9



Gambar 9. Frekuensi kata berdasarkan trigram untuk sentimen netral.

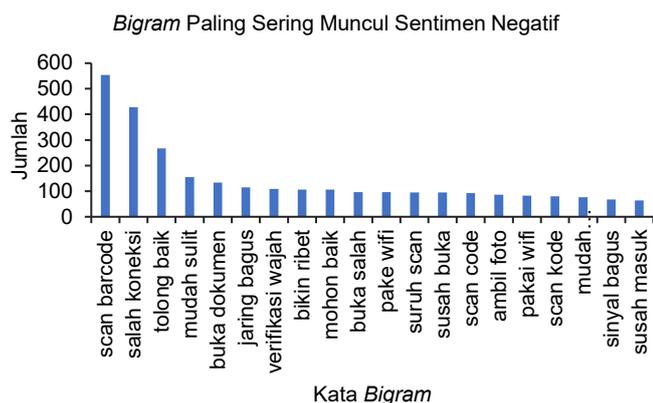


Gambar 10. Frekuensi kata berdasarkan unigram untuk sentimen negatif.

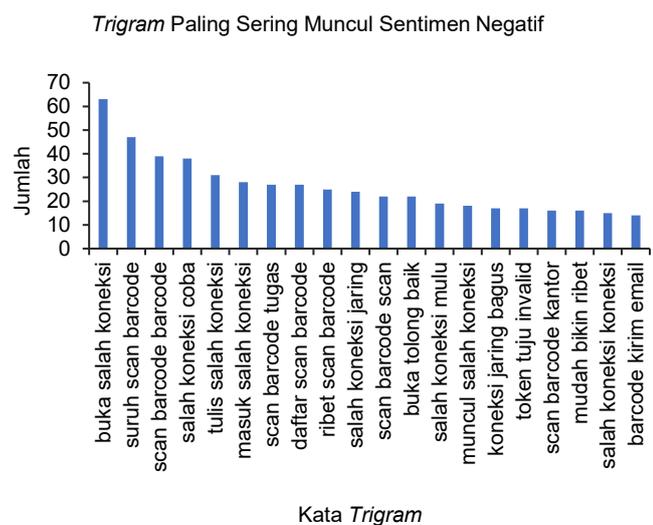
kemunculan) menekankan pentingnya koneksi internet yang stabil dalam penggunaan aplikasi.

Gambar 9 menunjukkan hasil implementasi trigram untuk kategori sentimen netral. Implementasi trigram pada sentimen netral menyoroti berbagai permasalahan teknis yang dialami oleh pengguna, khususnya terkait dengan konektivitas dan akses terhadap dokumen. Frasa-frasa seperti “scan barcode tugas” dan “suruh scan barcode” mencerminkan kebingungan yang mungkin dialami oleh sebagian pengguna dalam proses pemindaian barcode. Frasa seperti “salah koneksi coba” dan “buka salah koneksi” menunjukkan adanya kendala konektivitas yang menghambat kelancaran penggunaan aplikasi. Selain itu, trigram seperti “buka menu dokumen” dan “dokumen kartu keluarga” menggambarkan interaksi pengguna dengan dokumen penting, sedangkan frasa “tanda tangan elektronik” mengindikasikan prosedur verifikasi yang memerlukan perhatian lebih lanjut. Secara keseluruhan, hasil ini menekankan pentingnya peningkatan stabilitas teknis dan kemudahan akses pada aplikasi IKD.

Gambar 10 menunjukkan hasil implementasi unigram untuk kategori sentimen negatif. Implementasi unigram pada sentimen negatif menggambarkan berbagai keluhan pengguna yang berkaitan dengan aksesibilitas dan permasalahan teknis. Kata-kata seperti “buka” (2216 kemunculan), “scan” (1222 kemunculan), dan “barcode” (1050 kemunculan) sering muncul, mengindikasikan adanya permasalahan pada fitur pemindaian. Pengguna juga menyampaikan ketidakpuasan terhadap proses pendaftaran dan login melalui kata-kata seperti “daftar”, “data”, dan “masuk”. Sementara itu, kata-kata seperti



Gambar 11. Frekuensi kata berdasarkan bigram untuk sentimen negatif.



Gambar 12. Frekuensi kata berdasarkan trigram untuk sentimen negatif.

“ribet” (921 kemunculan) dan “susah” (848 kemunculan) mencerminkan tantangan yang dirasakan dalam menggunakan aplikasi. Permasalahan teknis lainnya, seperti “koneksi” (804 kemunculan) dan “salah” (746 kemunculan), serta permintaan perbaikan yang diungkapkan melalui kata “tolong” (740 kemunculan), menunjukkan tingkat frustrasi pengguna terhadap kinerja aplikasi IKD.

Gambar 11 menunjukkan hasil implementasi *bigram* untuk kategori sentimen negatif. Implementasi *bigram* pada sentimen negatif mengungkapkan keluhan utama pengguna terkait masalah pemindaian dan konektivitas. *Bigram* seperti “scan barcode” (554 kemunculan) menunjukkan kesulitan pengguna dalam menggunakan fitur pemindaian, yang juga didukung oleh frasa-frasa seperti “suruh scan” dan “scan code”. Permasalahan konektivitas tecermin melalui *bigram* seperti “salah koneksi” dan “buka dokumen” (masing-masing muncul sebanyak 428 dan 134 kali), meskipun beberapa pengguna menyatakan memiliki jaringan yang baik, sebagaimana terlihat dari frasa “jaring bagus” dan “sinyal bagus”. *Bigram* seperti “mudah sulit” dan “bikin ribet” menunjukkan adanya fitur yang dianggap membingungkan atau menyulitkan. Selain itu, pengguna juga menyampaikan harapan perbaikan melalui frasa seperti “tolong baik” dan “mohon baik”, yang mencerminkan keinginan akan peningkatan stabilitas dan kemudahan penggunaan aplikasi.

Gambar 12 menunjukkan hasil implementasi *trigram* untuk kategori sentimen negatif. Implementasi *trigram* pada sentimen negatif menyoroti permasalahan utama terkait konektivitas dan pemindaian *barcode*. *Trigram* seperti “buka salah koneksi” (63

kemunculan) dan “salah koneksi coba” (38 kemunculan) mencerminkan keluhan yang sering disampaikan pengguna mengenai gangguan jaringan. Selain itu, frasa seperti “tulis salah koneksi” dan “salah koneksi mulu” menunjukkan tingkat frustrasi pengguna akibat masalah koneksi yang berulang. Terkait dengan fitur pemindaian *barcode*, *trigram* seperti “suruh scan barcode” dan “ribet scan barcode” mengindikasikan kesulitan yang dihadapi dalam proses pemindaian. Di samping itu, *trigram* seperti “token tuju invalid” dan “barcode kirim email” mengarah pada permasalahan dalam proses autentikasi dan komunikasi data, yang turut memengaruhi pengalaman pengguna secara keseluruhan.

V. KESIMPULAN

Penelitian ini memanfaatkan 28.134 ulasan pengguna dari Google Play Store untuk mengklasifikasikan sentimen terhadap aplikasi IKD dengan menggunakan algoritma *random forest*. Model yang dibangun berhasil mencapai akurasi 83,85% dan *F1-score* sebesar 81,73% tanpa penerapan penyeimbangan kelas. Setelah penerapan metode SMOTE, nilai presisi mengalami peningkatan, meskipun akurasi dan *recall* sedikit menurun. Hal ini menunjukkan adanya *trade-off* antara keseimbangan data dan kemampuan generalisasi model.

Dibandingkan dengan penelitian terdahulu yang menggunakan algoritma *naïve Bayes*, SVM, dan KNN, model *random forest* menunjukkan hasil yang lebih konsisten dan kompetitif. Keberhasilan ini didukung oleh tahapan prapemrosesan yang menyeluruh serta ekstraksi fitur menggunakan metode TF-IDF. Model ini mampu menangani ketidakseimbangan kelas dan ragam ekspresi sentimen dalam skala data yang besar.

Selain itu, analisis *N-gram* berhasil mengungkap tema-tema umum dalam umpan balik pengguna, seperti kemudahan penggunaan serta keluhan terkait proses *login* dan konektivitas. Temuan ini menunjukkan bahwa pendekatan pembelajaran mesin tidak hanya bermanfaat dalam klasifikasi sentimen, tetapi juga mampu memberikan wawasan yang mendalam untuk mendukung perbaikan berbasis pengguna serta berkontribusi pada peningkatan kualitas layanan digital publik.

KONFLIK KEPENTINGAN

Para penulis menyatakan bahwa tidak terdapat konflik kepentingan dalam penelitian ini.

KONTRIBUSI PENULIS

Konseptualisasi, Kelvin H. dan Erlin; metodologi, Kelvin H. dan Erlin; perangkat lunak, Kelvin H.; validasi, Erlin; analisis formal, Kelvin H.; investigasi, Kelvin H.; sumber daya, Erlin, Yenny Desnelita, dan Dwi Oktarina; kurasi data, Kelvin H.; penulisan—draf awal, Kelvin H.; penulisan—tinjauan dan penyuntingan, Erlin, Yenny Desnelita, dan Dwi Oktarina; visualisasi, Kelvin H.; supervisi, Erlin; administrasi proyek, Erlin, Yenny Desnelita, dan Dwi Oktarina.

REFERENSI

- [1] R.W. Sasongko, “Implementasi identitas kependudukan digital di Kabupaten Bandung,” *J. Regist.*, vol. 5, no. 1, hal. 69–86, Sep. 2023, doi: 10.33701/jurnalregistratie.v5i1.3148.
- [2] P.C. Ardilia, “Optimalisasi pendampingan layanan administrasi kependudukan melalui program Kalimasada untuk mewujudkan tertib adminduk studi di Kelurahan Tembok Dukuh Kota Surabaya,” *PRAJA Obs., J. Penelit. Adm. Publik*, vol. 2, no. 3, hal. 63–68, Mei 2022.
- [3] A. Widiyarta dan I. Humaidah, “Implementasi aktivasi identitas kependudukan digital (IKD) dalam mendorong digitalisasi di Kelurahan Jepara Kota Surabaya,” *J. Ilm. Wahana Pendidik.*, vol. 9, no. 18, hal. 43–51, Sep. 2023, doi: 10.5281/zenodo.8310255.

- [4] M.N. Muttaqin dan I. Kharisudin, "Analisis sentimen pada ulasan aplikasi Gojek menggunakan metode support vector machine dan k nearest neighbor," *UNNES J. Math.*, vol. 10, no. 2, hal. 22–27, Nov. 2021, doi: 10.15294/ujm.v10i2.48474.
- [5] N.L.P.C. Savitri, R.A. Rahman, R. Venyutzky, dan N.A. Rakhmawati, "Analisis klasifikasi sentimen terhadap sekolah daring pada Twitter menggunakan supervised machine learning," *JuTISI (J. Tek. Inform. Sist. Inf.)*, vol. 7, no. 1, hal. 47–58, Apr. 2021, doi: 10.28932/jutisi.v7i1.3216.
- [6] A. Firdaus dan W.I. Firdaus, "Text mining dan pola algoritma dalam penyelesaian masalah informasi: (Sebuah ulasan)," *JUPITER, J. Penelit. Ilmu Teknol. Komput.*, vol. 13, no. 1, hal. 66–78, Apr. 2021.
- [7] A.C. Müller dan S. Guido, *Introduction to Machine Learning with Python: A Guide for Data Scientists*. Sebastopol, CA, AS: O'Reilly Media, 2017.
- [8] B.B. Baskoro, I. Susanto, dan S. Khomsah, "Analisis sentimen pelanggan hotel di Purwokerto menggunakan metode random forest dan TF-IDF (Studi kasus: Ulasan pelanggan pada situs TRIPADVISOR)," *J. Inform. Inf. Syst. Softw. Appl. (INISTA)*, vol. 3, no. 2, hal. 21–29, Mei 2021, doi: 10.20895/inista.v3i2.218.
- [9] A. Komarudin dan A.M. Hilda, "Analisis sentimen ulasan aplikasi identitas kependudukan digital pada Play Store menggunakan metode naïve Bayes," *Comput. Sci. (CO-SCIENCE)*, vol. 4, no. 1, hal. 28–36, Jan. 2024, doi: 10.31294/coscience.v4i1.2955.
- [10] P.K. Sari dan R.R. Suryono, "Komparasi algoritma support vector machine dan random forest untuk analisis sentimen metaverse," *J. Mnemon.*, vol. 7, no. 1, hal. 31–39, Feb. 2024, doi: 10.36040/mnemonic.v7i1.8977.
- [11] R.A. Lestari, A. Erfina, dan W. Jatmiko, "Penerapan algoritma support vector machine pada analisis sentimen terhadap identitas kependudukan digital," *J. Teknol. Inf. Ilmu Komput.*, vol. 10, no. 5, hal. 1063–1070, Okt. 2023, doi: 10.25126/jtiik.20231057264.
- [12] M. Ulfa, R.H. Kusumodestoni, dan A. Sucipto, "Analisis sentimen review aplikasi identitas kependudukan digital di Google Play Store menggunakan KNN," *J. Inform. Teknol. Sains (JINTEKS)*, vol. 6, no. 4, hal. 1155–1165, Nov. 2024, doi: 10.33558/jinteks.v6i4.4963.
- [13] A. Roihan, P.A. Sunarya, dan A.S. Rafika, "Pemanfaatan machine learning dalam berbagai bidang: Review paper," *IJCIT (Indones. J. Comput. Inf. Technol.)*, vol. 5, no. 1, hal. 75–82, Mei 2020, doi: 10.31294/ijcit.v5i1.7951.
- [14] R. Wati, S. Ernawati, dan H. Rachmi, "Pembobotan TF-IDF menggunakan naïve Bayes pada sentimen masyarakat mengenai isu kenaikan BIPIH," *J. Manaj. Inform. (JAMIKA)*, vol. 13, no. 1, hal. 84–93, Apr. 2023, doi: 10.34010/jamika.v13i1.9424.
- [15] Erlin dkk., "Dampak SMOTE terhadap kinerja random forest classifier berdasarkan data tidak seimbang," *Matrik: J. Manaj. Tek. Inform. Rekayasa Komput.*, vol. 21, no. 3, hal. 677–690, Jul. 2022, doi: 10.30812/matrik.v21i3.1726.
- [16] A. Mulianti, Y. Chrisnanto, dan H. Ashaury, "Optimalisasi klasifikasi support vector machine dengan SMOTE: Studi kasus ulasan pengguna aplikasi Alfagift," *J. Pekommas*, vol. 9, no. 2, hal. 249–258, Des. 2024, doi: 10.56873/jpkm.v9i2.5583.
- [17] R. Chairunnisa, Indriati, dan P.P. Adikara, "Analisis sentimen terhadap karyawan dirumahkan pada media sosial Twitter menggunakan fitur n-gram dan pembobotan augmented TF-IDF probability dengan k-nearest neighbour," *J. Pengemb. Teknol. Inf. Ilmu Komput.*, vol. 6, no. 4, hal. 1960–1965, Apr. 2022.
- [18] *Identitas Kependudukan Digital*. (2025). Kementerian Dalam Negeri. Diakses: 10 Maret 2025. [Online]. Tersedia: https://play.google.com/store/apps/details?id=gov.dukcapil.mobile_id&hl=id
- [19] E. Christy dan K. Suryowati, "Analisis klasifikasi status bekerja penduduk Daerah Istimewa Yogyakarta menggunakan metode random forest," *J. Stat. Ind. Komputasi*, vol. 6, no. 1, hal. 69–76, Jan. 2021.
- [20] D. Normawati dan S.A. Prayogi, "Implementasi naïve Bayes classifier dan confusion matrix pada analisis sentimen berbasis teks pada Twitter," *J-SAKTI (J. Sains Komput. Inform.)*, vol. 5, no. 2, hal. 697–711, Sep. 2021, doi: 10.30645/j-sakti.v5i2.369.