

Implementasi Algoritma Random Forest dalam Klasifikasi Ulasan Pengunjung Mall Semarang untuk Pengambilan Keputusan Layanan

Annisa Maizaliyanti¹, Khothibul Umam^{1,*}, Wenty Dwi Yuniarti¹, Maya Rini Handayani¹

¹ Program Studi Teknologi Informasi, Universitas Islam Negeri Walisongo Semarang, Indonesia

* Correspondence: khothibul_umam@walisongo.ac.id

Copyright: © 2025 by the authors

Received: 20 Mei 2025 | Revised: 1 Juni 2025 | Accepted: 3 Juli 2025 | Published: 12 Agustus 2025

Abstrak

Preferensi pengunjung terhadap *mall* di Semarang belum optimal karena ulasan daring belum dimanfaatkan secara maksimal dalam pengambilan keputusan. Penelitian kami bertujuan mengklasifikasikan sentimen ulasan *Google Maps* dari 13 *mall* di Semarang dengan total 2.600 ulasan. Pelabelan dilakukan secara manual berdasarkan rating, di mana rating 1–3 diasumsikan sebagai ulasan negatif dan 4–5 sebagai ulasan positif. Metode klasifikasi yang digunakan adalah *Random Forest* karena pendekatan *ensemble* (*bagging*) yang memberikan hasil optimal. Proses penelitian meliputi pengumpulan data, pelabelan, pembersihan, pembagian data, klasifikasi, dan evaluasi model. Data yang digunakan tidak seimbang dan didominasi oleh ulasan positif, sehingga diterapkan teknik *Synthetic Minority Over-sampling Technique* (SMOTE). Akurasi keseluruhan sebelum dan sesudah SMOTE tetap sama yaitu 84%. Meskipun demikian, performa model dalam mendeteksi ulasan negatif mengalami peningkatan *recall* semula 27% menjadi 44% dan *F1-score* dari 40% menjadi 52%, namun nilai tersebut masih tergolong rendah. Java Supermall Semarang menjadi *mall* dengan ulasan terbaik, dengan akurasi klasifikasi mencapai 90%. Model ini lebih baik mengenali ulasan positif, namun kurang andal untuk ulasan negatif. Oleh karena itu, penggunaannya sebagai preferensi pengambilan keputusan perlu dilakukan secara hati-hati. Penelitian ini membuka peluang pengembangan lebih lanjut, termasuk penggunaan model lain seperti *BERT* yang lebih unggul dalam memahami konteks dan bahasa dalam ulasan.

Kata kunci: *google maps; mall; preferensi; random forest; smote*

Abstract

Visitor preferences for malls in Semarang are not optimal because bold reviews have not been utilized optimally in decision making. Our research aims to classify the sentiment of Google Maps reviews from 13 malls in Semarang with a total of 2,600 reviews. Labeling is done manually based on ratings, where ratings 1–3 are considered negative reviews and 4–5 as positive reviews. The classification method used is Random Forest because the ensemble approach (*bagging*) provides optimal results. The research process includes data collection, labeling, cleaning, data sharing, classification, and model evaluation. The data used is unbalanced and dominated by positive reviews, so the Synthetic Minority Over-sampling Technique (SMOTE) technique was applied. The overall accuracy before and after SMOTE remained the same at 84%. However, the model's performance in detecting negative reviews increased from 27% to 44% in recall and *F1-score* from 0.40 to 0.52, but these values are still relatively low. Java Supermall Semarang is the mall with the best reviews, with a classification accuracy reaching 90%. This model is better at recognizing positive reviews, but less reliable for negative reviews. Therefore, its use as a decision-making preference needs to be done with caution. This research opens up opportunities for further development, including the use of other models such as *BERT* which are superior in understanding context and language in reviews.

Keywords: *google maps; mall; preference; random forests; smote*



PENDAHULUAN

Perkembangan teknologi yang semakin pesat, memudahkan masyarakat dalam mengakses berbagai informasi, termasuk mengenai ulasan publik seperti *mall* (Kadir et al., 2024). *Mall* sekarang menjadi bagian penting bagi masyarakat karena tidak hanya sebagai pusat perbelanjaan, tetapi sebagai tempat rekreasi, bersantai, dan tempat berkumpul (Adinegara et al., 2021). Desain *mall* saat ini dirancang sedemikian rupa untuk memberikan pengalaman belanja yang nyaman bagi pengunjung (Ranika et al., 2024). Sebelum mengunjungi suatu tempat, masyarakat cenderung mencari referensi terlebih dahulu melalui *platform* yang menyediakan ulasan *online* untuk memastikan kualitas dan fasilitas layanan yang ditawarkan. Sebagai kota metropolitan, Semarang memiliki banyak *mall* di mana *mall* tersebut terus berkembang sehingga persaingan untuk menarik minat pengunjung semakin ketat (Lu et al., 2021). Dalam hal ini, melalui data preferensi konsumen menjadi kunci untuk meningkatkan daya saing. Pemerintah juga turut mendorong peningkatan kualitas layanan publik melalui program *Mall Pelayanan Publik* di berbagai daerah sebagai upaya menghadirkan layanan yang efisien dan mudah diakses (Andriansyah, 2021). Data yang didapat adalah ulasan di *Google Maps* memberikan manfaat yang sangat penting karena mencerminkan sentimen publik secara langsung, autentik, dan jumlahnya besar. Analisis terhadap ulasan tersebut memungkinkan identifikasi kekuatan dan kelemahan layanan *mall*, serta mendukung pengambilan keputusan yang berbasis data. Hal ini menunjukkan bahwa pelayanan *mall* yang terarah untuk memenuhi kebutuhan pengunjung menjadi aspek penting dalam meningkatkan daya saing, terutama dalam penyediaan layanan barang dan jasa yang berkualitas (Aryanisila, 2023).

Banyaknya ulasan pengunjung yang tersedia di *platform* seperti *google maps* seringkali membuat pengelola *mall* mengalami kesulitan untuk menganalisis secara manual, terutama ketika jumlahnya mencapai ribuan dalam waktu singkat. Fitur ulasan di banyak situs juga sering tidak menyediakan pemisahan antara ulasan positif dan negatif, sehingga diperlukan analisis sentimen untuk mengklasifikasikannya (Aufan et al., 2024). Oleh karena itu, dibutuhkan teknologi yang secara otomatis bisa membaca dan memahami isi ulasan agar informasi penting bisa didapatkan dengan cepat dan akurat. Pada ulasan masyarakat dapat melihat penilaian, apakah bersifat positif atau negatif terhadap suatu tempat publik, seperti pusat perbelanjaan maupun tempat wisata (Khofifah et al., 2022). Ketidaksesuaian antara rating dan isi komentar seperti rating lima dengan komentar negatif bisa membingungkan calon pengunjung dan pengelola *mall*. Fenomena ini dikenal sebagai *Text-Rating Review Discrepancy* (TRRD), yakni ketidaksesuaian antara sentimen dalam teks ulasan dengan rating numerik yang diberikan (Almansour et al., 2022). Ketidaksesuaian ini tidak hanya membingungkan calon pengunjung, tetapi juga dapat menyesatkan pengelola dalam mengevaluasi kualitas layanan dan memahami kebutuhan pengunjung secara akurat.

Sebagai upaya untuk mengatasi pengelolaan ulasan pengunjung yang jumlahnya besar dan beragam, diperlukan pendekatan teknologi yang mampu mengolah informasi secara akurat dan efisien. Penelitian ini memanfaatkan teknologi *Natural Language Processing* (NLP). NLP adalah bidang kecerdasan buatan yang mempelajari interaksi manusia-komputer melalui bahasa alami (Furqan et al., 2023; Setiawan & Maharani, 2025), khususnya dalam subbidang *sentiment analysis* dan *text mining*, untuk mengklasifikasikan ulasan *Google Maps* ke dalam kategori positif dan negatif. Teknik ini memungkinkan pengelompokan data ulasan menjadi kategori positif dan negatif yang menggambarkan tingkat kepuasan pengguna terhadap layanan yang mereka terima (Rizki et al., 2023). *Natural Language Processing* (NLP) sendiri merupakan cabang ilmu yang mempelajari cara membuat komputer mampu memahami, memproses, dan menirukan cara berpikir serta berkomunikasi seperti manusia (Yahya et al., 2025). Salah satu pendekatan revolusioner dalam pemrosesan bahasa alami adalah penggunaan teknik *deep learning* untuk menghasilkan representasi kata yang kaya dan kontekstual (Adib et al., 2024). Metode yang digunakan adalah *Random Forest* karena mampu menangani data

teks yang memiliki variasi panjang, penggunaan bahasa informal, serta ketidakteraturan struktur kalimat yang umum dijumpai dalam ulasan *online*. Selain itu, *random forest* dipilih karena kemampuannya dalam mengurangi risiko *overfitting* yang sering terjadi pada model pohon keputusan tunggal, sehingga sangat sesuai untuk tugas klasifikasi teks dengan data yang tidak selalu terstruktur dengan baik (Mursyidah et al., 2024). Menurut Mamun et al. (2022), *random forest* cukup baik digunakan untuk klasifikasi teks, terutama jika digabung dengan TF-IDF. Algoritma *random forest* lebih akurat dibandingkan *support vector machine* dan *naive bayes* dengan peningkatan akurasi 7,16% (Fitri et al., 2020).

Penelitian sebelumnya yang dilakukan oleh Dista et al. (2023) menggunakan metode *k-means* yang dioptimasi dengan *particle swarm optimization* untuk mengelompokkan pengunjung *mall* berdasarkan data pendapatan dan pengeluaran. Hasil penelitian tersebut menunjukkan performa yang baik dengan ditandai oleh penurunan nilai *davies-bouldin index*, serta menyimpulkan bahwa pelanggan dengan pendapatan dan pengeluaran tinggi merupakan target utama pengelola *mall*. Namun, pendekatan ini hanya berfokus pada data kuantitatif dan tidak menangkap aspek kualitatif seperti kepuasan, opini, atau keluhan pengunjung yang dapat tercermin melalui ulasan. Padahal, ulasan seringkali memuat informasi preferensi dan sentimen yang tidak terlihat dalam data numerik. Oleh karena itu, penelitian ini berupaya mengisi kesenjangan tersebut dengan menganalisis ulasan pengunjung menggunakan pendekatan berbasis teks dan *sentiment analysis*, untuk memperoleh pemahaman yang lebih utuh terhadap pengalaman dan persepsi pengunjung *mall*.

Penelitian Panjaitan & Supriadi (2023) melakukan analisis sentimen terhadap ulasan pengunjung wisata Heritage di Kota Semarang, yaitu Lawang Sewu, Museum Kota Lama, dan Sam Poo Kong, menggunakan metode klasifikasi *Naive Bayes* dengan bantuan perangkat lunak *RapidMiner*. Ulasan diambil dari *google maps* sebanyak 50 ulasan terbaru per lokasi, dan diklasifikasikan menjadi dua kategori yaitu positif dan negatif. Hasil akurasi tertinggi diperoleh dari Museum Kota Lama sebesar 91%, sementara Lawang Sewu dan Sam Poo Kong sama-sama memperoleh akurasi 68,5%. Pemilihan dua kelas sentimen dilakukan untuk meningkatkan akurasi dan menghindari ambiguitas dalam klasifikasi.

Berbeda dengan penelitian sebelumnya, penelitian kami ini fokus menganalisis sentimen pengunjung *mall* di Semarang menggunakan klasifikasi dua kelas, yaitu positif dan negatif, tanpa melibatkan kategori netral yang seringkali ambigu. Penelitian Dista et al. (2023) hanya menyoroti data numerik seperti pendapatan dan pengeluaran, sedangkan Panjaitan & Supriadi (2023) menganalisis sentimen ulasan pada wisata Heritage, bukan *mall*. Belum ditemukan studi sebelumnya yang secara khusus meneliti preferensi atau ulasan pengunjung terhadap *mall* di kota Semarang, sehingga penelitian ini mengisi kekosongan tersebut. Dengan tujuan penelitian mengklasifikasikan sentimen ulasan *google maps* dari 13 *mall* di Semarang. Diharapkan hasil klasifikasi ini, pengunjung bisa melihat *mall* dengan rating tertinggi sebagai referensi tambahan dalam pengambilan keputusan.

METODE

Penelitian ini termasuk dalam jenis penelitian kuantitatif dengan menggunakan metode *random forest* untuk melakukan klasifikasi preferensi pengunjung *mall* di Semarang. Desain penelitian ini mengikuti tahapan klasifikasi berbasis *machine learning*, yaitu pengumpulan data, *pre-processing* data, pembagian data latih dan data uji, pelabelan data, pembangunan model menggunakan algoritma *random forest*, dan evaluasi performa model. Dataset di peroleh menggunakan teknik secara manual satu per satu dari *platform google maps* dengan menggunakan filter ulasan terbaru. Data yang dikumpulkan disimpan dalam *file* dengan format *CSV* untuk kemudian dilakukan pengolahan data di tahap *pre-processing*. Masing-masing *mall* diambil 200 ulasan sehingga data yang diperoleh dari 13 *mall* yaitu 2.600 dataset.



Gambar 1. Alur penelitian

Proses *pre-processing* mencakup pembersihan data, *case folding* (mengubah huruf kapital menjadi kecil), *tokenizing* (memisah kata), *stopword removal* (menghapus kata umum), dan *stemming* (mengubah kata ke bentuk dasar) menggunakan *library* seperti *Pandas*, *NLTK*, *Sastrawi*, dan *Emoji* (Apriliyani et al., 2024). Seluruh ulasan telah dikonversi ke dalam Bahasa Indonesia. Representasi teks dilakukan menggunakan metode TF-IDF untuk mengubah teks menjadi fitur numerik. Kemudian data dipecah dengan rasio 80:20 menjadi data latih data uji (Cahyani & Budiman, 2025). Pelabelan data dilakukan secara otomatis berdasarkan rating pengguna di *Google Maps*. Rating 1 sampai 3 dikategorikan sebagai sentimen negatif, sedangkan rating 4 dan 5 dikategorikan sebagai sentimen positif. Untuk mengatasi ketidakseimbangan jumlah data antar kelas sentimen, digunakan metode SMOTE (Synthetic Minority Over-sampling Technique). SMOTE berfungsi untuk menyeimbangkan data latih dengan cara membuat sampel sintesis dari kelas minoritas, sehingga model tidak bias terhadap kelas mayoritas dan dapat meningkatkan performa klasifikasi.

Random forest merupakan metode turunan yang berasal dari algoritma *decision tree*. Masing-masing pohon keputusan dalam *Random Forest* dilatih menggunakan sampel yang dipilih acak dan hanya sebagian data yang digunakan. Dalam penelitian ini, digunakan sebanyak 100 pohon keputusan dengan parameter *random state* 42 untuk memastikan hasil yang konsisten. Algoritma *random forest* menggunakan pendekatan yang dikenal sebagai metode *bagging*, di mana setiap pohon keputusan diberi subset fitur yang berbeda dari data asli. Proses pengacakan fitur ini, juga dikenal sebagai *bootstrapping*, membantu dalam menghasilkan model yang lebih stabil dan akurat. Keunggulan lainnya, *random forest* mampu mengurangi risiko *overfitting* dengan cara memilih fitur berdasarkan kriteria tertentu sehingga mengurangi ketergantungan antar fitur. Oleh karena itu, *random forest* dapat menghasilkan hasil klasifikasi yang baik. Keputusan kelas dalam klasifikasi *random forest* ditentukan oleh pohon keputusan dengan suara terbanyak. Secara matematis proses ini dapat dirumuskan menggunakan persamaan 1.

$$F(x) = \frac{1}{j} \sum_{j=1}^J h_j(x) \tag{1}$$

Keterangan :

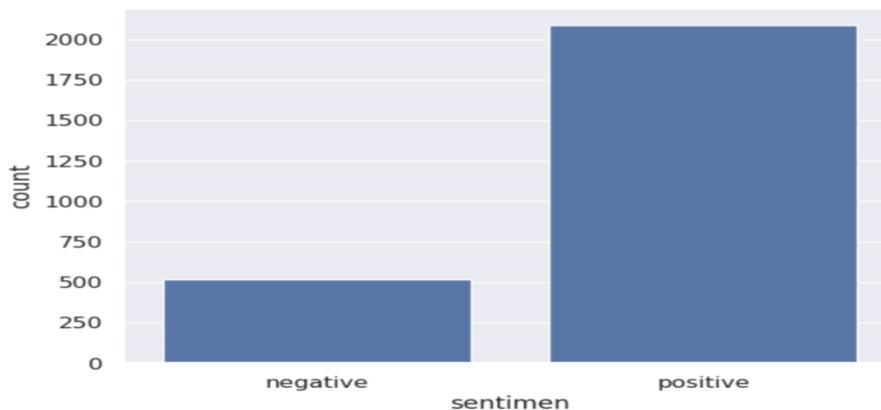
- $F(x)$: output
- J : jumlah pohon dalam *ensemble*
- $h_j(x)$: output dari pohon ke-(j)

Persamaan 1, $F(x)$ diperoleh dari rata-rata *output* seluruh pohon keputusan untuk menghasilkan prediksi yang akurat. Evaluasi dilakukan pada data uji menggunakan matrik akurasi, *precision*, *recall*, *f1-score*, dan *support*. Evaluasi ini bertujuan mengukur kemampuan model dalam mengklasifikasikan data baru. Selain itu, *word cloud* dari ulasan positif dan negatif digunakan untuk mengidentifikasi kata-kata yang sering muncul, membantu pengelola *mall* dalam mengevaluasi pelayanan.

HASIL DAN PEMBAHASAN

Hasil

Proses pengumpulan data dilakukan secara manual satu per satu dari *platform Google Maps* dengan menggunakan filter ulasan terbaru. Masing-masing *mall* diambil sebanyak 200 ulasan terbaru dari 13 *mall* yang ada di Kota Semarang, sehingga total dataset yang diperoleh sebanyak 2.600 data. Penelitian ini menggunakan dua kelas sentimen, yaitu positif dan negatif. Dari total data tersebut, terdapat 2.080 ulasan berlabel positif dan 520 ulasan berlabel negatif. Oleh karena itu, digunakan algoritma *Synthetic Minority Over-sampling Technique (SMOTE)* untuk menyeimbangkan jumlah data pada masing-masing kelas. Distribusi data awal dari masing-masing kelas sentimen dapat dilihat pada gambar 2.



Gambar 2. Data awal

Tahap selanjutnya adalah pengujian pada data uji. Tujuannya untuk melihat sejauh mana model bisa mengklasifikasikan data baru secara akurat. Dengan mengevaluasi model, bisa diketahui apakah model *overfitting* (sesuai data pelatihan) atau *underfitting* (kurang belajar). Pada penelitian ini evaluasi dilakukan hanya sekali pada data uji, tanpa menggunakan teknik validasi silang, sehingga hasil evaluasi hanya merepresentasikan performa model pada subset data tersebut. Evaluasi model fokus pada data uji karena matrik data latih sering terlalu optimistis akibat *overfitting*. Meskipun data uji bisa bias, hasilnya lebih mencerminkan kemampuan model pada data baru. Hasil pelatihan model dapat dilihat pada tabel 1.

Tabel 1. Evaluasi model dengan akurasi 84%

| | <i>Precision</i> | <i>Recall</i> | <i>F1-score</i> | <i>Support</i> |
|---------------------|------------------|---------------|-----------------|----------------|
| <i>Negative</i> | 78% | 27% | 40% | 103 |
| <i>Positive</i> | 85% | 98% | 91% | 417 |
| <i>Accuracy</i> | | | 84% | 520 |
| <i>Macro avg</i> | 81% | 63% | 66% | 520 |
| <i>Weighted avg</i> | 83% | 84% | 81% | 520 |

Berdasarkan hasil pada tabel 1 menunjukkan model menghasilkan akurasi 84% dari total data yang diuji berhasil diklasifikasikan dengan benar. Pada kelas positif, *precision* 85% prediksi benar. Nilai *Recall* 98% berhasil di kenali, dan *F1-score* 91% merupakan gabungan dari *precision* dan *recall* menunjukkan performa yang sangat baik. Sedangkan kelas negatif *precision* 78 % prediksi negatif benar. *Recall* hanya 27% data negatif yang dikenali dengan benar. *F1-score* 40% performa model dalam mengenali kelas negatif masih rendah. Dari hasil tersebut model masih cenderung bias ke kelas mayoritas, oleh karena itu dilakukan teknik

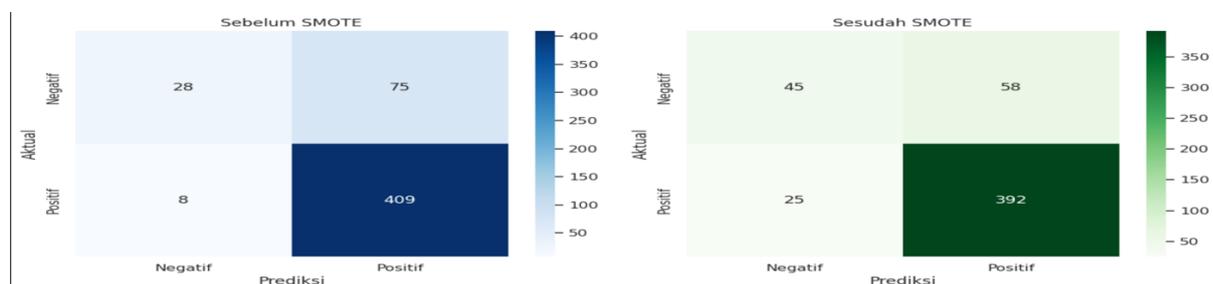
SMOTE untuk menyeimbangkan data pelatihan agar model bisa mengenali kelas minoritas dengan baik. Hasil SMOTE dapat dilihat pada tabel 2.

Tabel 2. Evaluasi model dengan *smote* (akurasi 84%)

| | <i>Precision</i> | <i>Recall</i> | <i>F1-score</i> | <i>Support</i> |
|---------------------|------------------|---------------|-----------------|----------------|
| <i>Negative</i> | 64% | 44% | 52% | 103 |
| <i>Positive</i> | 87% | 94% | 90% | 417 |
| <i>Accuracy</i> | | | 84% | 520 |
| <i>Macro avg</i> | 76% | 69% | 71% | 520 |
| <i>Weighted avg</i> | 83% | 84% | 83% | 520 |

Berdasarkan hasil SMOTE pada tabel 2 menunjukkan model menghasilkan akurasi 84% dari total data yang diuji berhasil diklasifikasikan dengan benar. Pada kelas positif, *precision* 87% prediksi benar. Nilai *Recall* 94% berhasil di kenali, dan *F1-score* 90% menunjukkan performa yang sangat baik. Sedangkan kelas negatif *precision* 64% prediksi negatif benar. *Recall* mengalami kenaikan dengan nilai 44% data negatif yang dikenali dengan benar. *F1-score* 52% performa model dalam mengenali kelas negatif. Meskipun terdapat peningkatan setelah dilakukan SMOTE, namun performa pada kelas negatif masih tergolong rendah. Rendahnya *recall* pada kelas negatif salah satu penyebabnya adalah jumlah data negatif yang terlalu sedikit, sehingga meskipun sudah dilakukan *oversampling* dengan SMOTE, model masih belum bisa mengenali pola karakteristik data negatif dengan baik.

Berdasarkan gambar 3 menunjukkan kinerja model sebelum dan sesudah diterapkannya teknik SMOTE. Setelah diterapkannya SMOTE, model menjadi lebih baik dalam mengenali kelas positif dan negatif. Sebelum SMOTE, model cenderung bias terhadap kelas positif, ditunjukkan oleh tingginya *false positive* dengan nilai 75 dan rendahnya *true negative* dengan nilai 28. Setelah SMOTE, jumlah *true negative* meningkat 45 dan *false positive* menurun 58, meskipun terjadi sedikit peningkatan *false negative* dari 8 menjadi 25. Secara keseluruhan, SMOTE membantu mengurangi bias dan meningkatkan kemampuan model dalam mendeteksi kedua kelas dengan lebih adil.



Gambar 3. confusion matrik sebelum dan sesudah *smote*

Word cloud merupakan visualisasi kata yang muncul dalam kumpulan teks, dan sangat bermanfaat untuk eksplorasi awal data. Gambar 4 menunjukkan *word cloud* dari komentar berlabel positif dengan kata-kata seperti baik, bagus, ok, dan lumayan yang mencerminkan kepuasan pengunjung. Sementara itu, gambar 5 menampilkan *word cloud* negatif dengan kata-kata seperti parkir, toilet, jorok, susah, dan kotor, yang mencerminkan keluhan terutama pada aspek kebersihan dan kenyamanan. Namun *word cloud* bersifat statistik dan tidak mempertimbangkan konteks atau kontribusi kata terhadap hasil prediksi model.

Berdasarkan gambar 6 menunjukkan hasil *f1-score* per kelas untuk model klasifikasi yang membedakan positif dan negatif. *F1-score* merupakan matrik evaluasi yang menggabungkan nilai *precision* dan *recall*, yang bisa mengukur performa model terutama pada

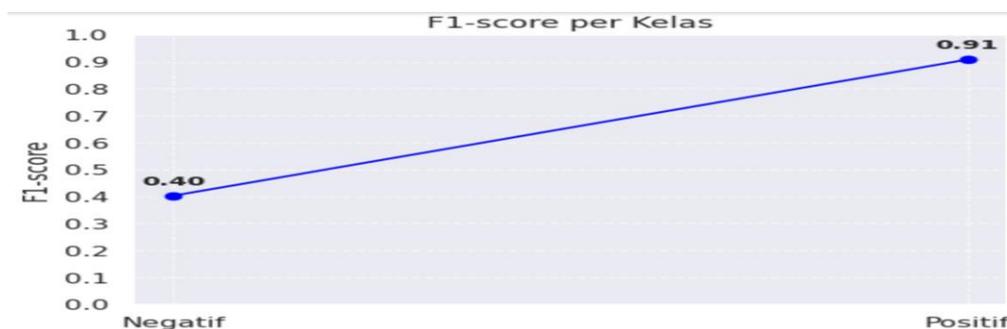
data yang tidak seimbang. Dalam grafik ini, terlihat bahwa model memiliki *f1-score* sebesar 40% untuk kelas negatif dan 91% untuk kelas positif. Perbedaan yang cukup signifikan ini menunjukkan bahwa model jauh lebih baik dalam mengenali kelas positif dibandingkan kelas negatif. Rendahnya *F1-score* pada kelas negatif bisa menjadi indikasi bahwa model sering melakukan kesalahan dalam mengklasifikasikan data negatif, baik berupa *false positive* maupun *false negative*. Ketimpangan performa ini disebabkan oleh data yang tidak seimbang sehingga model yang terlalu bias terhadap kelas positif.



Gambar 4. Word cloud positif



Gambar 5. Word cloud negatif



Gambar 6. F1-score setiap kelas

Pembahasan

Hasil penelitian dari model *random forest* menunjukkan kinerja yang baik dalam menganalisis ulasan pengunjung *mall*, dengan akurasi keseluruhan sebesar 84%. Akurasi tersebut menunjukkan bahwa model mampu dalam memprediksi sentimen positif dan negatif dari komentar *google maps*. Namun ketidakseimbangan dataset yang digunakan dapat menyebabkan bias model, di mana model lebih cenderung memprediksi ulasan sebagai positif, sehingga potensi kesalahan dalam mengidentifikasi ulasan negatif meningkat. Oleh karena itu, akurasi saja tidak cukup untuk menilai performa model secara menyeluruh.

Pemilihan model *random forest* ini didasarkan pada kemampuannya untuk menangani data dengan berbagai fitur serta mengurangi risiko *overfitting* melalui metode *ensemble learning*. Model dilatih menggunakan data yang sudah melalui tahap pelabelan dan pemrosesan. Pada penelitian ini dataset yang digunakan tidak seimbang, jadi model cenderung bias terhadap kelas mayoritas. Ketidakseimbangan data ini menjadi tantangan utama pada penelitian ini, dimana jumlah komentar positif sebanyak 2080 data jauh lebih banyak dibandingkan komentar negatif yang hanya 520 data. Untuk menyeimbangkan data tersebut dilakukan proses SMOTE. Tujuan dilakukan SMOTE untuk meningkatkan kemampuan model dalam mengenali sentimen yang kompleks, terutama kelas negatif yang sebelumnya sulit untuk dikenali.

Setelah dilakukan teknik SMOTE, akurasi keseluruhan tidak mengalami peningkatan signifikan. Peningkatan yang jelas setelah dilakukan SMOTE yaitu pada matrik *recall* untuk kelas negatif yaitu dari 27% menjadi 44%. Hal ini menunjukkan bahwa SMOTE efektif dalam menyeimbangkan data dan membantu model mengenali pola-pola pada ulasan negatif yang sebelumnya kurang terdeteksi dengan baik.

Faktor yang sangat mendukung penelitian ini yaitu pemilihan dataset dari *Google Maps* yang relevan dan bisa mempresentasikan pengalaman nyata pengunjung saat ke *mall*. Ulasan tersebut mencakup berbagai aspek penting seperti kenyamanan, kelengkapan tenant, dan pelayanan. Keragaman tersebut memungkinkan model untuk mengenali sentimen positif dan negatif termasuk mengenali pola kata yang berkaitan dengan aspek spesifik seperti kenyamanan atau keluhan fasilitas. Namun, perlu diketahui bahwa tidak semua komentar di *Google Maps* menggunakan Bahasa Indonesia, sehingga ulasan tersebut perlu diterjemahkan terlebih dahulu menggunakan fitur terjemahan di *google maps*.

Sebagai perbandingan dengan penelitian sebelumnya oleh Dista et al. (2023) menggunakan *k-means* yang dioptimasi dengan *particle swarm optimization* untuk mengelompokkan pengunjung *mall* berdasarkan data pendapatan dan pengeluaran, dengan hasil baik ditandai dengan penurunan *davies-bouldin index*. Namun, pendekatan ini fokus pada data kuantitatif tanpa mempertimbangkan opini atau keluhan dalam ulasan. Sementara itu penelitian yang dilakukan oleh Panjaitan & Supriadi (2023) menggunakan metode *naive bayes* untuk analisis sentimen ulasan wisata Heritage di Semarang dengan dua kelas yaitu positif dan negatif, yang menghasilkan akurasi tertinggi 91% pada Museum Kota Lama, serta 68,5% pada Lawang Sewu dan Sam Poo Kong. Pendekatan dua kelas ini dinilai efektif untuk meningkatkan akurasi klasifikasi.

Penelitian ini melengkapi kekurangan dua penelitian sebelumnya dengan menggabungkan pendekatan berbasis teks menggunakan algoritma *random forest* dan klasifikasi dua kelas sentimen yaitu positif dan negatif. Dengan memfokuskan analisis pada ulasan sebagai data kualitatif, kita bisa mendapatkan gambaran yang lebih menyeluruh tentang persepsi dan pengalaman pengunjung *mall*. Secara praktis, model ini berpotensi membantu pengunjung dalam memperoleh gambaran yang lebih jelas mengenai pengalaman pengunjung lain di *mall*, mencakup aspek kenyamanan, kelengkapan tenant, dan pelayanan. Dengan demikian, pengguna dapat membuat keputusan yang lebih tepat dan memuaskan sesuai preferensi mereka. Namun, pengguna dan pengembang aplikasi juga harus menyadari keterbatasan model dan terus berupaya memperbaiki ketepatan klasifikasi, terutama pada sentimen negatif yang lebih kompleks.

SIMPULAN

Pada penelitian ini menunjukkan bahwa *Random Forest* efektif menganalisis ulasan pengunjung *mall* Semarang dengan akurasi 84% setelah penerapan SMOTE. Hasil ini menunjukkan model cukup baik mengenali sentimen positif dengan *f1-score* 90%, namun dalam mengenali sentimen negatif masih kurang baik dengan *f1-score* 52% akibat data yang tidak seimbang. Penerapan SMOTE berhasil mengurangi bias, karena peningkatan *true negative* dan penurunan *false positive*, meskipun *false negative* naik. Kontribusi utama penelitian ini adalah menunjukkan bahwa data yang seimbang dan tahap pemrosesan yang baik dapat menghasilkan model yang andal dengan data terbatas. Secara praktis, model ini bermanfaat untuk mendukung pengambilan keputusan bagi pengunjung dalam memilih *mall*. Selanjutnya, penelitian dapat diperdalam dengan memanfaatkan model berbasis *transformer* seperti *BERT* untuk lebih memahami aspek yang mempengaruhi sentimen.

REFERENSI

- Adib, K., Handayani, M. R., Yuniarti, W. D., & Umam, K. (2024). Opini Publik Pasca-Pemilihan Presiden: Eksplorasi Analisis Sentimen Media Sosial X Menggunakan SVM. *SINTECH Journal*, 7(2), 80–91. <https://doi.org/10.31598/sintechjournal.v7i2.1581>
- Almansour, N. M. (2022). Triple-negative breast cancer: A Brief Review About Epidemiology, Risk Factors, Signaling Pathways, Treatment and Role of Artificial Intelligence. *Frontiers in Molecular Biosciences*, 9(1), 1–15. <https://doi.org/10.3389/fmolb.2022.836417>
- Andriansyah, A., Sulastri, E., & Satispi, E. (2021). The Role of Government Policies in Environmental Management. *Research Horizon*, 1(3), 86–93. <https://doi.org/10.54518/rh.1.3.2021.86-93>
- Apriliyani, M., Musyaffaq, M. I., Nur'Aini, S., Handayani, M. R., & Umam, K. (2024). Implementasi Analisis Sentimen Pada Ulasan Aplikasi Duolingo Di Google Playstore Menggunakan Algoritma Naïve Bayes. *AITI: Jurnal Teknologi Informasi*, 21(2), 298–311. <https://doi.org/10.24246/aiti.v21i2.298-311>
- Aryanisila, A. (2023). Penyelenggaraan Mall Pelayanan Publik untuk Masyarakat di Kota Palembang. *SIMBOL (Jurnal Administrasi Publik dan Pemerintahan)*, 2(1), 1–11. <https://doi.org/10.55850/simbol.v2i1.65>
- Aufan, M. H., Handayani, M. R., Nurjanna, A. B., Wibowo, N. C. H., & Umam, K. (2024). The Perceptions Of Semarang Five Star Hotel Tourists with Support Vector Machine On Google Reviews. *Jurnal Teknik Informatika (JUTIF)*, 5(5), 1241–1247. <https://doi.org/10.52436/1.jutif.2024.5.5.2025>
- Cahyani, O. N., & Budiman, F. (2025). Performa Logistic Regression dan Naive Bayes dalam Klasifikasi Berita Hoax di Indonesia. *Edumatic: Jurnal Pendidikan Informatika*, 9(1), 60–68. <https://doi.org/10.29408/edumatic.v9i1.28987>
- Dista, T. M., & Abdulloh, F. F. (2022). Clustering Pengunjung Mall Menggunakan Metode K-Means Dan Particle Swarm Optimization. *Jurnal Media Informatika Budidarma*, 6(3), 1339–1348. <https://doi.org/10.30865/mib.v6i3.4172>
- Fitri, E., Yuliani, Y., & Gata, W. (2020). Analisis Sentimen Terhadap Aplikasi Ruangguru Menggunakan Algoritma Naive Bayes, Random Forest, dan Support Vector Machine. *TRANSFORMATIKA*, 1(1), 71–80.
- Furqan, M., Sriani, & Shidqi, M. N. (2023). Chatbot Telegram Menggunakan Natural Language Processing. *Walisongo Journal of Information Technology*, 5(1), 15-26. <https://doi.org/10.21580/wjit.2023.5.1.14793>
- Kadir, I., Kenanga, D. T., & Kasetty, D. M. (2024). Pengaruh Online Customer Review Dan Online Customer Rating Terhadap Keputusan Pembelian Produk Di Online Marketplace Shopee (Studi Kasus Mahasiswa STIE Petra Bitung). *Tangkoko Jurnal Akuntansi dan Manajemen*, 10(1), 41–50. <https://doi.org/10.33795/jab.v10i1.3601>
- Khofifah, W., Rahayu, D. N., & Yusuf, A. M. (2022). Analisis Sentimen Menggunakan Naive Bayes Untuk Melihat Review Masyarakat Terhadap Tempat Wisata Pantai Di Kabupaten Karawang Pada Ulasan Google Maps. *Jurnal Interkom: Jurnal Publikasi Ilmiah Bidang Teknologi Informasi dan Komunikasi*, 16(4), 171-180. <https://doi.org/10.35969/interkom.v16i4.192>
- Lu, C., Kuswoyo, C., Abednego, F., & Josephine, S. G. (2021). Pengaruh Faktor Lingkungan Dan Pengalaman Belanja Mall Terhadap Perilaku Belanja Mall. *Jurnal Inspirasi Bisnis dan Manajemen*, 5(1), 87–100. <https://doi.org/10.33603/jibm.v5i1.4937>
- Mamun, A. A., Islam, M. S., & Hasan, M. K. (2022). A Novel Improved Random Forest For Text Classification Using Feature Selection And TF-IDF. *Journal of King Saud University – Computer and Information Sciences*, 34(8), 5667–5674. <https://doi.org/10.1016/j.jksuci.2022.06.001>

- Mursyidah, M., Davi, M., & Novitri, S. D. (2024). Klasifikasi Sentimen Review Pengguna terhadap Aplikasi Instagram menggunakan Algoritma Random Forest. *Jurnal Infomedia: Teknik Informatika, Multimedia, dan Jaringan*, 9(2), 106–115.
- Panjaitan, C. H. P., & Supriadi, C. (2023). Analisis Sentimen Pengunjung Wisata Heritage Kota Semarang Menggunakan Naive Bayes Pada Ulasan Google Maps. *JETI: Jurnal Elektronika dan Teknologi Informasi*, 4(2), 1–28.
- Ranika, V. G., Mutrofin, A., & Nathania, E. C. A. (2024). Analisis Preferensi Mall Culture Dalam Gaya Hidup Konsumtif. *WISSEN: Jurnal Ilmu Sosial dan Humaniora*, 2(3), 89–96. <https://doi.org/10.62383/wissen.v2i3.175>
- Rizki, A. F., Prihartono, W., & Fathurrohman. (2023). Analisis Sentimen Ulasan Google Maps Rumah Sakit Khalishah Di Cirebon Dengan Algoritma Naive Bayes. *JITET (Jurnal Informatika dan Teknik Elektro Terapan)*, 13(2), 728–738. <https://doi.org/10.23960/jitet.v13i2.6309>
- Setiawan, A. D. H., & Maharani, W. (2025). Understanding Public Sentiments on the 2024 Presidential Election through BERT-Powered Analysis. *Edumatic: Jurnal Pendidikan Informatika*, 9(1), 89-98. <https://doi.org/10.29408/edumatic.v9i1.29267>
- Yahya, F., Ratnawati, D. E., & Rahayudi, B. (2025). Analisis Sentimen Ulasan Pengguna dari Google Maps Menggunakan Metode Long Short-Term Memory (Studi Kasus: Rumah Sakit Gatoel). *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 9(4), 1–11.