

KLASIFIKASI SENTIMEN WISATAWAN MUSEUM GEDUNG JUANG 45 BEKASI BERDASARKAN ULASAN GOOGLE MAPS DENGAN ALGORITMA KNN

Ilham Rakan Zaffran, Rina Candra Noor Santi

Teknik Informatika, Universitas Stikubank Semarang
Jalan Raya Tri Lomba Juang 1 Semarang, Indonesia
ilhamrakanzaffran@mhs.unisbank.ac.id

ABSTRAK

Perkembangan teknologi digital memungkinkan pemanfaatan ulasan pengguna pada platform seperti Google Maps sebagai sumber data dalam memahami persepsi publik terhadap destinasi wisata. Namun, banyaknya data ulasan yang bersifat tidak terstruktur menjadi tantangan dalam menganalisis opini secara manual. Penelitian ini bertujuan untuk mengklasifikasikan sentimen ulasan wisatawan terhadap Museum Gedung Juang 45 Bekasi secara otomatis dengan pendekatan text mining. Metode yang digunakan meliputi pengumpulan data ulasan melalui SerpAPI, dilanjutkan dengan tahapan preprocessing (cleaning, case folding, normalisasi, stemming, stopword removal), dan pembobotan fitur menggunakan Term Frequency-Inverse Document Frequency (TF-IDF). Proses klasifikasi dilakukan dengan algoritma K-Nearest Neighbor (KNN). Hasil klasifikasi menunjukkan bahwa 67% ulasan tergolong positif, 22% negatif, dan 11% netral, dengan akurasi model mencapai 74% dan skor F1 rata-rata 0,52. Temuan ini menunjukkan bahwa KNN cukup efektif dalam mengidentifikasi persepsi publik, serta dapat digunakan sebagai dasar pengambilan keputusan untuk peningkatan kualitas layanan museum.

Kata kunci : Analisis Sentimen, K-Nearest Neighbor, Google Maps, Museum Gedung Juang, Text Mining

1. PENDAHULUAN

Perkembangan teknologi digital telah merubah drastis cara masyarakat berinteraksi dan berbagi pengalaman, khususnya melalui platform berbasis lokasi seperti Google Maps. Platform ini menyediakan fitur ulasan teks dan penilaian bintang yang memberikan wawasan langsung tentang kepuasan pengunjung terhadap destinasi wisata [1]. Ulasan pengguna tersebut mencerminkan opini publik yang masif dan realtime, dan menjadi sumber data penting untuk pengambilan keputusan berbasis data [2].

Museum Gedung Juang 45 Bekasi adalah bangunan bersejarah peninggalan kolonial Belanda dan situs perjuangan rakyat Bekasi yang mulai dibangun pada tahun sekitar 1906 oleh Khouw Tjeng Kie [3]. Museum ini saat ini menjadi salah satu destinasi edukatif penting, yang juga menarik cukup banyak ulasan di Google Maps [4].

Namun, volume dan sifat tidak terstruktur dari ulasan tersebut menyulitkan pengelolaan manual. Oleh karena itu, dibutuhkan pendekatan otomatis melalui **analisis sentimen**, yaitu teknik text mining untuk mengidentifikasi opini pengguna dalam kategori positif, negatif, maupun netral [5]. Teknik ini telah terbukti efektif dalam sektor pariwisata, seperti di kawasan Borobudur dan Prambanan yang menggunakan analisis aspek berbasis Google Maps [6]

Dalam konteks klasifikasi, algoritma sederhana seperti K-Nearest Neighbor (KNN) banyak digunakan karena kemudahannya dan performanya yang kompetitif pada dataset ulasan wisata [7]. Selain itu, teknik TF-IDF umum dipakai sebagai representasi vektor fitur teks dan telah menunjukkan akurasi tinggi dalam klasifikasi sentimen ulasan wisata [8]

Penelitian ini bertujuan untuk mengklasifikasikan sentimen ulasan wisatawan terhadap Museum Gedung Juang 45 Bekasi yang diperoleh dari Google Maps, menggunakan kombinasi persiapan data melalui TF-IDF dan klasifikasi dengan algoritma KNN[9]. Tujuannya adalah menghasilkan pemetaan opini publik yang mendalam dan akurat, sehingga membantu pengelola museum dalam memahami persepsi pengunjung serta meningkatkan kualitas layanan berbasis data.

2. TINJAUAN PUSTAKA

Penelitian terkait analisis sentimen dalam sektor pariwisata telah menunjukkan kemajuan signifikan seiring berkembangnya teknologi pengolahan teks. Berbagai studi terdahulu menunjukkan bahwa algoritma *K-Nearest Neighbor* (KNN) efektif digunakan untuk mengklasifikasikan opini pengunjung berdasarkan ulasan digital[5]Berhasil menerapkan KNN dalam mengelompokkan ulasan pengguna Shopee dan mencapai akurasi sebesar 88%, memperkuat validitas pendekatan ini dalam skenario opini berbasis teks. Hasil serupa juga ditemukan oleh Nugroho dan Wibowo[8]yang menggunakan kombinasi KNN dan Decision Tree untuk mengolah ulasan wisata budaya, mencapai akurasi 86% dengan teknik *ensemble voting*. Penerapan KNN dalam konteks pariwisata terbukti mampu mengklasifikasikan sentimen positif, negatif, dan netral secara akurat [10]. Salah satu tantangan utama dalam analisis sentimen adalah ketidakseimbangan distribusi data antar kategori sentimen, khususnya sentimen netral yang sering sulit dikenali. Untuk mengatasi hal ini, sejumlah penelitian menekankan pentingnya tahap *text preprocessing* yang meliputi

normalisasi, stemming, dan *stopword removal*, yang dapat meningkatkan kualitas data input dan akurasi model[11] menunjukkan bahwa tahapan preprocessing yang baik dapat memperkuat klasifikasi model KNN, bahkan ketika data ulasan berasal dari berbagai platform dengan gaya bahasa yang berbeda. Selain itu, penggunaan teknik pembobotan fitur seperti TF-IDF juga dinilai krusial dalam[12]

Dalam konteks pariwisata digital, metode *text mining* telah menjadi tulang punggung dalam mengolah opini pengunjung yang tidak terstruktur. Teknik ini tidak hanya memungkinkan ekstraksi informasi dari teks ulasan, tetapi juga mendukung pengambilan keputusan berbasis data oleh pengelola destinasi wisata[13]. Visualisasi data melalui wordcloud turut memperkuat proses eksplorasi informasi, memungkinkan pengelola untuk mengidentifikasi kata-kata kunci yang sering muncul dalam ulasan, baik positif maupun negatif [14]

Penelitian ini memberikan dasar teoretis yang mendukung pemilihan metode K-Nearest Neighbor (KNN) dalam analisis sentimen ulasan Museum Gedung Juang 45 Bekasi. Dengan memanfaatkan teknik text preprocessing seperti tokenisasi, normalisasi, dan stemming, serta evaluasi performa model menggunakan confusion matrix, penelitian ini diharapkan dapat menghasilkan klasifikasi sentimen yang akurat[15]. Hasil analisis ini bermanfaat untuk memahami persepsi pengunjung secara lebih menyeluruh dan memberikan masukan strategis bagi pengelola dalam meningkatkan kualitas layanan dan pengalaman wisata di destinasi tersebut.

3. METODE PENELITIAN

Penelitian ini bertujuan untuk menganalisis sentimen ulasan pengunjung Museum Gedung Juang 45 Bekasi menggunakan metode *K-Nearest Neighbor* (KNN). Penelitian dilakukan dalam lima tahapan utama, yaitu pengumpulan data, preprocessing data, pelabelan sentimen, pemodelan klasifikasi, dan evaluasi performa model.

3.1. Pengumpulan Data

Data diperoleh secara otomatis melalui teknik *web crawling* menggunakan SerpAPI untuk mengambil ulasan dari Google Maps. Data yang dikumpulkan terdiri dari nama pengguna, isi ulasan, rating bintang, dan waktu ulasan. Data disimpan dalam format .CSV untuk keperluan analisis lanjutan menggunakan bahasa pemrograman Python.

3.2. Preprocessing Data

Tahapan preprocessing bertujuan untuk membersihkan dan menyiapkan data teks agar dapat dianalisis secara efektif. Proses preprocessing meliputi:

Cleaning: Menghapus karakter tidak penting seperti angka, simbol, URL, emoji, dan tanda baca.

- Case Folding: Mengubah semua huruf menjadi huruf kecil untuk menghindari perbedaan makna karena kapitalisasi.
- Normalisasi : Mengubah kata tidak baku menjadi bentuk baku, misalnya "nggak" menjadi "tidak".
- Stemming: Mengembalikan kata ke bentuk dasarnya, seperti "bermain" menjadi "main", menggunakan algoritma. Teknik ini menggunakan library NLP seperti Sastrawi untuk bahasa Indonesia
- Stopword Removal: Menghapus kata-kata umum seperti "dan", "yang", "itu", yang tidak membawa makna penting.

3.3. Pelabelan Data

Pelabelan sentimen dilakukan oleh pengguna dengan mengacu pada nilai rating dari Google Maps:

- Rating 4–5 → Sentimen Positif
- Rating 3 → Sentimen Netral
- Rating 1–2 → Sentimen Negatif

Label ini kemudian dikonfirmasi secara manual berdasarkan konteks isi ulasan untuk memastikan kesesuaian.

3.4. Pemodelan Klasifikasi

Algoritma K-Nearest Neighbor (KNN) adalah metode klasifikasi berbasis instance yang digunakan dalam pembelajaran terawasi (*supervised learning*). KNN bekerja dengan prinsip bahwa data yang memiliki karakteristik serupa akan berada pada jarak yang berdekatan dalam ruang fitur. Ketika sebuah data baru akan diklasifikasikan, KNN akan menghitung jarak antara data tersebut dengan seluruh data latih, kemudian mengambil k tetangga terdekat untuk menentukan mayoritas kelas dari tetangga-tetangga tersebut sebagai prediksi kelas data baru.

Dalam penelitian ini, setiap ulasan yang telah direpresentasikan dalam bentuk vektor melalui metode TF-IDF akan dibandingkan menggunakan jarak Euclidean. Nilai parameter k diuji dalam beberapa variasi untuk mendapatkan performa terbaik model klasifikasi. Rumus Euclidean Distance yang digunakan adalah sebagai berikut:

$$d(x, y) = \sqrt{\sum_{i=1}^n (X_i - Y_i)^2} \quad (1)$$

3.5. Evaluasi Model

Evaluasi dilakukan menggunakan Confusion Matrix dengan empat metrik utama, yaitu Akurasi, Precision, Recall, dan F1-Score, untuk masing-masing kategori sentimen.

- Accuracy: Mengukur seberapa banyak prediksi model yang benar dari keseluruhan data uji.

$$\text{Accuracy} = \frac{\text{Jumlah prediksi benar}}{\text{Total data}} = \frac{TP_{pos} + TP_{neg} + TP_{net}}{\text{Total Data}} \quad (2)$$

- b. Precision: Mengukur ketepatan model dalam memprediksi suatu kelas.

$$PrecisionPos = \frac{TPpos}{TPpos + FPpos}$$

$$PrecisionNeg = \frac{TPneg}{TPneg + FPneg}$$

$$PrecisionNet = \frac{TPnet}{TPnet + FPnet}$$

(3)

- c. Recall: Mengukur kemampuan model dalam menemukan semua data yang benar untuk suatu kelas.

$$RecallPos = \frac{TPpos}{Total\ data\ positif}$$

$$RecallNeg = \frac{TPneg}{Total\ data\ negatif}$$

$$RecallNet = \frac{TPnet}{Total\ data\ netral}$$

(4)

- d. F1-Score: Rata-rata harmonis antara precision dan recall, digunakan untuk menilai kinerja model secara keseluruhan

$$F1Pos = 2 \times \frac{Precisionpos \times Recallpos}{Precisionpos + Recallpos}$$

$$F1Neg = 2 \times \frac{Precisionneg \times Recallneg}{Precisionneg + Recallneg}$$

$$F1Net = 2 \times \frac{Precisionnet \times Recallnet}{Precisionnet + Recallnet}$$

(5)

3.6. Hasil Evaluasi Model

Hasil analisis menunjukkan bahwa distribusi sentimen ulasan terhadap Museum Gedung Juang 45 Bekasi terdiri atas 67% positif, 22% negatif, dan 11% netral. Evaluasi dilakukan untuk mengukur performa algoritma *K-Nearest Neighbor* (KNN) dalam mengklasifikasikan sentimen menggunakan data uji. Metrik evaluasi mencakup precision, recall, F1-score, dan akurasi. Berdasarkan pengujian, model menghasilkan akurasi sebesar 0,74 atau 74% prediksi sesuai label sebenarnya, yang menunjukkan kinerja cukup baik dalam memetakan opini pengunjung secara umum.

Kendala utama ditemukan pada klasifikasi sentimen netral, yang sering memiliki kemiripan pola bahasa dengan ulasan positif maupun negatif, sehingga menurunkan nilai evaluasi pada kelas tersebut. Meskipun demikian, model tetap memberikan hasil yang relevan sebagai dasar pemahaman persepsi publik dan dapat menjadi bahan evaluasi dalam peningkatan kualitas layanan wisata secara digital.

4. HASIL DAN PEMBAHASAN

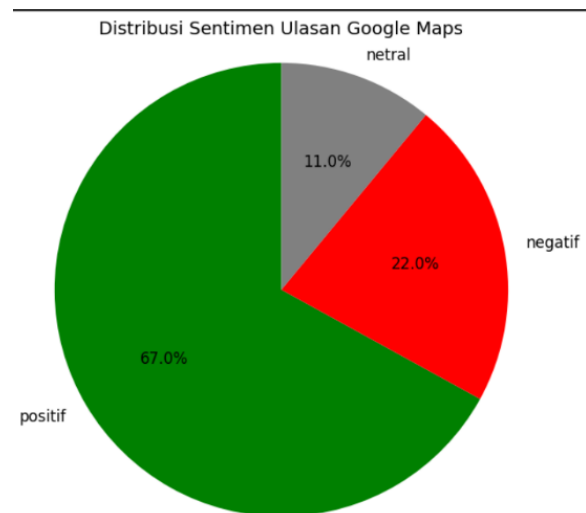
4.1. Hasil Pengumpulan Data dan Praproses Data

Penelitian ini menggunakan data ulasan pengunjung Museum Gedung Juang 45 Bekasi yang diambil dari Google Maps dengan teknik *web scraping* menggunakan SerpAPI. Dari proses ini, diperoleh 567 data ulasan mentah yang mencakup nama pengguna,

rating bintang, tanggal ulasan, dan isi teks. Data tersebut kemudian diproses melalui tahapan cleaning, case folding, stemming, dan stopword removal untuk membersihkan dan menyiapkan teks. Setelah proses praproses selesai, tersisa 500 data ulasan bersih yang siap digunakan dalam proses klasifikasi sentimen.

- Ulasan positif (rating 4–5) sebanyak 67%
- Ulasan netral (rating 3) sebanyak 11%
- Ulasan negatif (rating 1–2) sebanyak 22%

Distribusi ini divisualisasikan dalam diagram pada Gambar 1.



Gambar 1. Distribusi Sentimen Ulasan Pengunjung Museum Gedung Juang 45

4.2. Klasifikasi Sentimen Menggunakan KNN

Setelah data praproses selesai, dilakukan klasifikasi sentimen menggunakan algoritma *K-Nearest Neighbor* (KNN). Algoritma ini mengklasifikasikan data berdasarkan kedekatan jarak antar vektor ulasan, dengan menghitung jarak Euclidean antar data. Data yang telah dibobotkan dengan TF-IDF digunakan sebagai input model. Penentuan nilai parameter k dilakukan secara eksperimen, dan nilai terbaik diperoleh pada $k = 5$. Model menghasilkan klasifikasi ke dalam tiga kategori sentimen, yaitu positif, negatif, dan netral, berdasarkan analisis terhadap vektor fitur dari masing-masing ulasan.

4.3. Evaluasi Kinerja Model

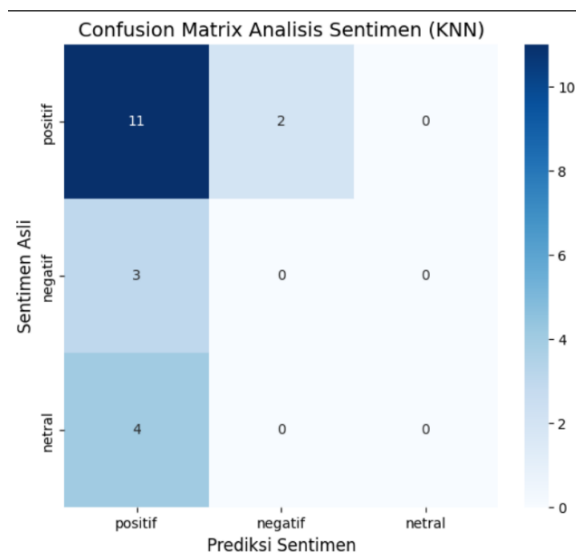
Evaluasi dilakukan untuk mengukur performa algoritma *K-Nearest Neighbor* (KNN) dalam mengklasifikasikan sentimen ulasan. Evaluasi ini mencakup penggunaan confusion matrix dan perhitungan metrik akurasi, precision, recall, serta F1-score.

Berdasarkan hasil pengujian menggunakan data uji, diperoleh hasil sebagai berikut:

Tabel 1 . Tabel Model

Matrix	Nilai
Accuracy	74%
Precision	0,74
Recall	0,74
F1-Score	0,52

Meskipun akurasi cukup tinggi, nilai F1-score yang lebih rendah menunjukkan bahwa model masih mengalami kesulitan, khususnya dalam mengklasifikasikan sentimen netral. Hal ini disebabkan oleh kemiripan pola bahasa netral dengan kategori positif dan negatif, serta ketidakseimbangan jumlah data antar kelas.



Gambar 2. Confisuion Marix

Pada Gambar 2 Confusion Matrix menggambarkan perbandingan antara nilai aktual dan hasil prediksi dalam analisis sentimen ulasan Museum Gedung Juang 45 Bekasi menggunakan algoritma K-Nearest Neighbor (KNN). Dalam matriks ini, baris merepresentasikan jumlah data berdasarkan kategori sentimen yang sebenarnya, sedangkan kolom menunjukkan hasil prediksi dari model. Nilai-nilai yang terletak pada diagonal matriks mencerminkan prediksi yang sesuai dengan sentimen asli, sementara nilai-nilai di luar diagonal mengindikasikan kesalahan klasifikasi. Berdasarkan hasil visualisasi, model KNN berhasil mengklasifikasikan 11 ulasan positif secara tepat, namun gagal sepenuhnya dalam mengenali sentimen negatif dan netral, karena semua data dari kedua kategori tersebut justru diprediksi sebagai sentimen positif. Hal ini menunjukkan bahwa model mengalami kesulitan dalam membedakan pola bahasa dari sentimen negatif dan netral yang mungkin memiliki kemiripan dengan bahasa positif. Dengan demikian, evaluasi ini menunjukkan bahwa meskipun model memiliki performa cukup baik dalam mendeteksi sentimen positif, perlu adanya perbaikan lebih lanjut terutama pada pemrosesan data dan penyeimbangan kelas agar akurasi klasifikasi untuk

semua kategori sentimen dapat ditingkatkan secara merata.

4.4. Analisis Wordcloud

Untuk memahami kata-kata dominan yang muncul dalam tiap kategori sentimen, digunakan visualisasi *wordcloud*. Wordcloud dibuat untuk masing-masing kelas sentimen: positif, negatif, dan netral. Hasil visualisasi menunjukkan bahwa kata-kata seperti "bagus", "sejarah", "menarik", dan "keluarga" sering muncul dalam sentimen positif, sedangkan kata-kata seperti "ramai", "parkir", "kotor", dan "panas" mendominasi sentimen negatif. Untuk sentimen netral, kata yang muncul cenderung bersifat deskriptif dan tidak mengandung muatan emosi yang jelas. Visualisasi ini membantu dalam memahami fokus perhatian dan pengalaman pengunjung dari perspektif yang lebih intuitif.



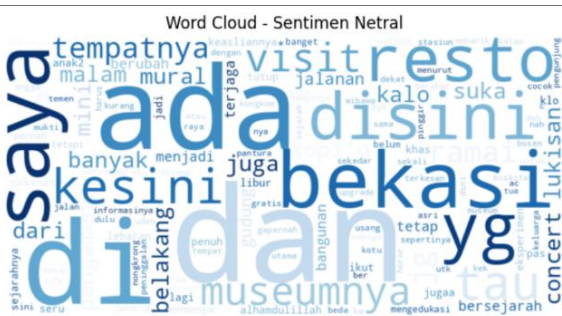
Gambar 3. Wordcloud Positif

Gambar 3 menunjukkan word cloud yang terbentuk dari ulasan-ulasan yang diklasifikasikan sebagai sentimen positif. Kata-kata yang paling sering muncul dalam kategori ini antara lain: "ada", "dan", "juga", "museum", "bagus", "tempat", "enak", "nyaman", "ramah", dan "luas". Kemunculan kata "bagus" dan "nyaman" mengindikasikan bahwa para pengunjung merasa puas dengan pengalaman mereka, baik dari sisi kondisi tempat maupun suasana yang dirasakan selama berkunjung. Kata "ramah" mengarah pada interaksi positif dengan petugas atau staf museum, yang dianggap memberikan pelayanan yang menyenangkan. Sementara itu, kata "luas" dan "tempat" menggambarkan bahwa pengunjung merasa area museum cukup memadai dan tidak sempit, yang dapat menambah kenyamanan selama berwisata. Selain itu, kata-kata seperti "juga", "untuk", dan "ada" merupakan kata hubung atau kata bantu yang memang sering muncul dalam kalimat, namun dalam konteks ini, mereka turut berkontribusi pada struktur narasi positif dari para pengulas. Word cloud ini secara umum mencerminkan bahwa sentimen positif didominasi oleh kata-kata yang bernuansa menyenangkan, menghargai, dan puas terhadap fasilitas serta pengalaman mengunjungi museum.



Gambar 4. Wordcloud Negatif

Gambar 4 merepresentasikan word cloud dari ulasan-ulasan yang masuk dalam kategori sentimen negatif. Kata-kata dominan yang terlihat antara lain: "bekasi", "dan", "yang", "masuk", "tambun", "parkir", "gedung", "luar", "dulu", dan "kelelawar". Ulasan dengan sentimen negatif ini banyak mengandung keluhan yang berkaitan dengan akses masuk, kondisi fisik gedung, hingga lingkungan sekitar museum. Misalnya, kata "masuk" dan "parkir" sering dikaitkan dengan keluhan mengenai tiket masuk yang mungkin dianggap mahal atau tempat parkir yang kurang memadai. Kata "gedung", "luar", dan "dulu" menunjukkan kemungkinan adanya kritik terhadap kondisi bangunan yang dinilai kurang terawat atau terlalu tua. Bahkan munculnya kata "kelelawar" dapat mengindikasikan adanya keluhan terhadap kebersihan atau gangguan dari hewan di dalam museum. Sementara itu, kata "tambun", "wilayah", dan "selatan" yang berkaitan dengan lokasi dapat menjadi indikator bahwa sebagian pengunjung mengeluhkan aksesibilitas atau letak museum yang kurang strategis. Word cloud ini menggambarkan bahwa ulasan negatif cenderung fokus pada aspek fisik dan logistik dari museum yang belum memuaskan pengunjung.



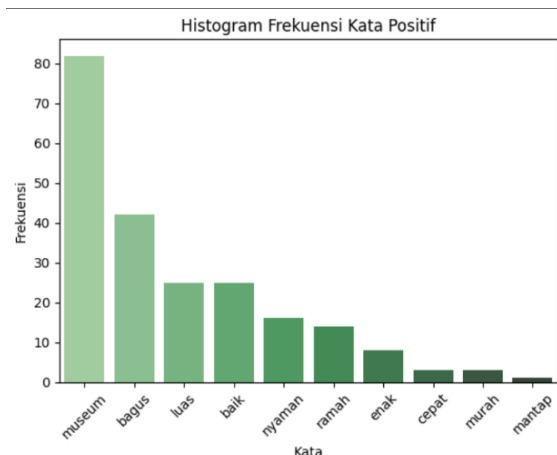
Gambar 5. Wordcloud Netral

Gambar 5 menunjukkan word cloud dari ulasan yang dikategorikan sebagai sentimen netral. Beberapa kata utama yang muncul adalah: "saya", "ada", "disini", "bekasi", "resto", "visit", "lukisan", "mural", "gratis", dan "kesini". Berbeda dengan sentimen positif atau negatif, sentimen netral cenderung berisi narasi yang bersifat informatif, deskriptif, atau narasi pribadi tanpa muatan emosional yang kuat. Misalnya, kata "saya", "disini", dan "kesini" menunjukkan kalimat berbentuk pernyataan tentang kunjungan atau

pengalaman pribadi tanpa penilaian tertentu. Kata "resto", "lukisan", dan "mural" menunjukkan bahwa pengunjung mendeskripsikan hal-hal yang mereka lihat atau temukan di sekitar museum, tanpa menyatakan apakah mereka menyukai atau tidak. Kemunculan kata "gratis" kemungkinan menginformasikan bahwa tidak dikenakan biaya untuk masuk, namun tidak disampaikan dalam nada pujian maupun keluhan. Word cloud ini memberikan gambaran bahwa sentimen netral berisi observasi dan penjelasan faktual dari pengunjung tanpa konotasi emosional yang mencolok.

4.5. Histogram Distribusi Sentimen Diagram Batang

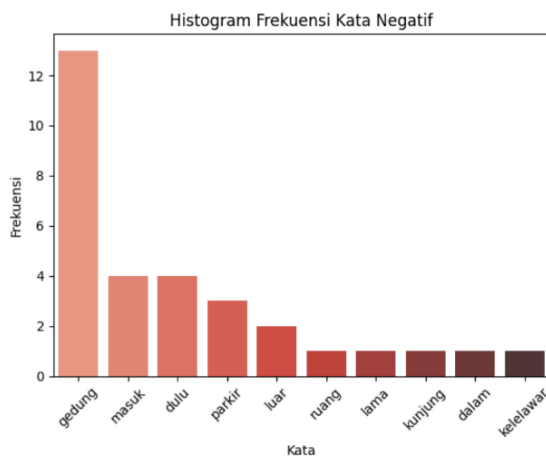
Analisis distribusi sentimen dilakukan dengan menggunakan histogram berupa *diagram batang*, yang menggambarkan jumlah ulasan dalam masing-masing kategori sentimen: positif, negatif, dan netral. Visualisasi ini bertujuan untuk memperjelas persebaran opini pengunjung terhadap Museum Gedung Juang 45 Bekasi. Berdasarkan hasil klasifikasi, diketahui bahwa jumlah ulasan positif mendominasi, yaitu sebesar 67%, disusul oleh 22% ulasan negatif, dan 11% ulasan netral. Diagram batang menunjukkan bahwa mayoritas pengunjung memberikan tanggapan yang baik terhadap museum ini, baik dari segi fasilitas, nilai sejarah, maupun pengalaman kunjungan secara keseluruhan. Tingginya proporsi sentimen positif dalam histogram menunjukkan bahwa citra Museum Gedung Juang 45 di mata pengunjung cukup baik. Sementara itu, keberadaan sentimen negatif dan netral dalam jumlah yang lebih kecil tetap perlu diperhatikan oleh pengelola sebagai bahan evaluasi untuk peningkatan layanan. Visualisasi dalam bentuk histogram ini membantu menyampaikan informasi secara lebih intuitif dan cepat dipahami oleh pembaca.



Gambar 6. Diagram Batang Positif

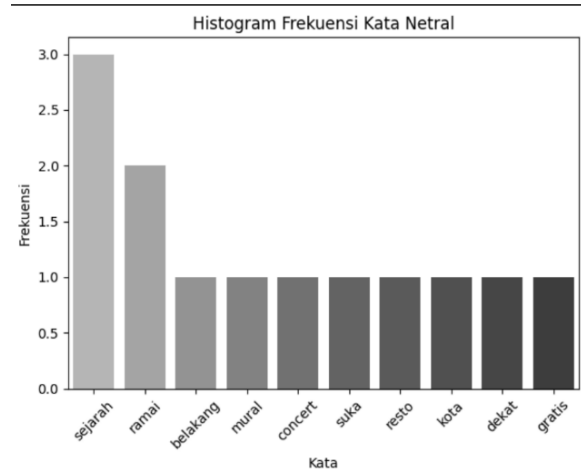
Gambar 6 adalah histogram yang menampilkan kata-kata yang paling sering muncul dalam ulasan dengan sentimen positif. Kata "museum" menempati posisi teratas dengan frekuensi sangat tinggi, yaitu

lebih dari 80 kali, menunjukkan bahwa pengunjung yang merasa puas sering menyebutkan kata ini secara langsung dalam konteks positif. Selanjutnya kata "bagus" muncul sebanyak lebih dari 40 kali, yang jelas mencerminkan bahwa banyak pengunjung menyatakan kepuasan terhadap fasilitas atau pengalaman mereka. Kata-kata lain yang juga signifikan dalam kategori ini adalah "luas", "baik", "nyaman", "ramah", dan "enak", yang menandakan kenyamanan lingkungan, keramahan pelayanan, hingga kesan menyenangkan secara keseluruhan. Kata-kata seperti "cepat", "murah", dan "mantap" juga memperkuat nuansa positif dari ulasan-ulasan tersebut. Histogram ini menunjukkan bahwa pengunjung yang memiliki pengalaman menyenangkan cenderung memuji secara eksplisit kondisi tempat, pelayanan, serta nilai yang mereka rasakan selama berada di museum.



Gambar 7. Diagram Batang Negatif

Gambar 7 menggambarkan histogram untuk kata-kata yang paling sering muncul dalam ulasan berisi sentimen negatif. Kata yang paling menonjol adalah "gedung" dengan frekuensi tertinggi mencapai 13 kali, menunjukkan adanya keluhan yang cukup sering terkait dengan kondisi bangunan, seperti kerusakan, kurang terawat, atau tidak sesuai ekspektasi. Disusul oleh kata "masuk" dan "dulu" yang masing-masing muncul sebanyak 4 kali, mengindikasikan adanya keluhan terhadap proses masuk (mungkin terkait harga tiket atau antrean) dan nostalgia masa lalu yang membandingkan kondisi sekarang yang dianggap menurun. Kata "parkir" muncul sebanyak 3 kali, yang dapat mencerminkan keluhan tentang ketersediaan atau akses area parkir. Sementara kata-kata seperti "luar", "ruang", "lama", "kunjung", "dalam", dan "kelelawar" muncul lebih sedikit, namun tetap signifikan. Secara khusus, kata "kelelawar" menunjukkan keluhan unik yang bisa terkait dengan kebersihan atau gangguan hewan di area dalam gedung. Histogram ini menunjukkan bahwa sentimen negatif banyak mengarah pada kritik terhadap aspek fisik dan fungsional dari museum.



Gambar 8. Diagram Batang Netral

Gambar 8 merupakan histogram untuk kata-kata yang termasuk dalam kategori sentimen netral, yang umumnya tidak menyiratkan emosi positif atau negatif secara langsung, melainkan deskriptif atau informatif. Kata "sejarah" menjadi yang paling sering muncul dengan frekuensi 3 kali, menandakan bahwa pengunjung menyebutkan tema sejarah dalam narasi mereka, tanpa memberikan penilaian tertentu. Kata "ramai" muncul sebanyak 2 kali, yang dapat mencerminkan observasi kondisi keramaian pengunjung tanpa menyatakan apakah itu hal yang menyenangkan atau tidak. Kata-kata lain seperti "belakang", "mural", "concert", "suka", "resto", "kota", "dekat", dan "gratis" masing-masing muncul sebanyak 1 kali. Kata-kata ini lebih banyak memberikan gambaran mengenai lokasi, acara, atau suasana yang dijumpai di sekitar museum. Misalnya, kata "mural" dan "concert" bisa mengacu pada elemen visual atau aktivitas yang berlangsung di sana, sementara "gratis" dan "dekat" memberikan informasi praktis tentang biaya dan aksesibilitas. Histogram ini menunjukkan bahwa dalam kategori netral, ulasan cenderung berbentuk pernyataan observatif dan tidak menekankan pada emosi atau penilaian subjektif.

4.6. Pembahasan

Hasil penelitian menunjukkan bahwa mayoritas ulasan pengunjung terhadap Museum Gedung Juang 45 Bekasi memiliki sentimen positif, dengan persentase mencapai 67%. Hal ini menunjukkan bahwa sebagian besar wisatawan merasa puas dengan pengalaman kunjungan mereka, khususnya terhadap aspek nilai sejarah bangunan, suasana lingkungan, serta keramahan petugas di lokasi. Sementara itu, ulasan dengan sentimen negatif tercatat sebesar 22%, yang umumnya berisi kritik terhadap kurangnya fasilitas pendukung, ketidaknyamanan area parkir, hingga kondisi kebersihan. Adapun sentimen netral sebanyak 11%, mencerminkan adanya ulasan yang bersifat deskriptif atau informatif tanpa mengandung ekspresi emosional yang kuat.

Evaluasi performa model menggunakan confusion matrix menunjukkan bahwa algoritma K-

Nearest Neighbor (KNN) memiliki tantangan dalam mengklasifikasikan sentimen netral, karena sering kali dikategorikan secara keliru sebagai sentimen positif atau negatif. Hal ini terjadi akibat kemiripan pola bahasa dalam ulasan netral dengan dua kategori lainnya, yang membuat model kesulitan dalam membedakan secara eksplisit. Nilai akurasi model yang diperoleh sebesar 74%, dengan F1-score rata-rata sebesar 0,60, menunjukkan bahwa kinerja klasifikasi tergolong sedang, dan masih dapat ditingkatkan dengan pengayaan data atau penerapan teknik balancing antar kelas. Berdasarkan hasil analisis *wordcloud*, ulasan positif didominasi oleh kata-kata seperti "bagus", "bersejarah", dan "menarik". Sementara itu, ulasan negatif sering memuat kata-kata seperti "panas", "kurang rapi", dan "macet", yang mencerminkan ketidaknyamanan pengunjung terhadap beberapa aspek fisik dan aksesibilitas. Sedangkan kata-kata dalam ulasan netral cenderung bersifat umum, seperti "tempat", "museum", atau "lokasi", tanpa menyiratkan ekspresi suka atau tidak suka secara jelas.

Secara keseluruhan, hasil penelitian ini memberikan gambaran menyeluruh mengenai persepsi pengunjung terhadap Museum Gedung Juang 45 Bekasi melalui analisis sentimen digital, serta menunjukkan bahwa pendekatan analisis berbasis algoritma KNN dapat menjadi metode yang relevan untuk membantu pengelola destinasi dalam melakukan evaluasi dan pengambilan keputusan berbasis data.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian, dapat disimpulkan bahwa mayoritas ulasan pengunjung terhadap Museum Gedung Juang 45 Bekasi memiliki sentimen positif sebesar 67%, menunjukkan bahwa secara umum pengunjung merasa puas dengan pengalaman wisata mereka. Algoritma K-Nearest Neighbor (KNN) yang digunakan dalam klasifikasi sentimen menunjukkan performa yang cukup baik dengan akurasi sebesar 74% dan F1-score rata-rata 0,60, meskipun masih terdapat tantangan dalam mengklasifikasikan sentimen netral akibat kemiripan pola bahasa. Untuk penelitian selanjutnya, disarankan agar jumlah data, khususnya pada kelas netral, diperbanyak dan algoritma pembandingan seperti Naïve Bayes atau SVM turut dieksplorasi guna memperoleh hasil klasifikasi yang lebih optimal dan seimbang.

DAFTAR PUSTAKA

- [1] E. N. Hamdana, A. O. Nur Wardani, and A. R. Tri Hayati Ririd, "Sentiment Analysis of Visitor Reviews on Google Maps at Kampung Coklat Tourism," *Journal of Artificial Intelligence and Software Engineering (J-AISE)*, vol. 5, no. 1, p. 274, Mar. 2025, doi: 10.30811/jaise.v5i1.6488.
- [2] S. Widodo and B. Hartono, "Analisis Sentimen Pengguna Google Terhadap Destinasi Wisata Di Kota Semarang Menggunakan Metode K-Nearest Neighbor," *Jurnal Ilmiah Komputer*, vol. 2, no. 2, pp. 545–554, Aug. 2023.
- [3] I. Imbron, M. Falah, and N. Kartika, "Revitalisasi Bangunan Cagar Budaya Gedung Juang Tambun sebagai Museum Digital Kabupaten Bekasi, 2019-2021," *Fajar Historia: Jurnal Ilmu Sejarah dan Pendidikan*, vol. 8, no. 1, pp. 110–128, Apr. 2024, doi: 10.29408/fhs.v8i1.25308.
- [4] M. N. Kholifah and W. Nurjayanti, "PENGARUH REVITALISASI BANGUNAN CAGAR BUDAYA GEDUNG JUANG 45 TERHADAP KEAKTIFAN BERWISATA SEJARAH SEBAGAI MUSEUM DIGITALISASI PERTAMA DI JABODETABEK," *Jurnal Ilmiah Arsitektur*, vol. 3, no. 3, pp. 11–19, Mar. 2022, [Online]. Available: <http://siar.ums.ac.id/>
- [5] H. Sibyan and N. Hasanah, "ANALISIS SENTIMEN PADA WISATA DIENG DENGAN ALGORITMA K-NEAREST NEIGHBOR (K-NN)," *Jurnal Penelitian dan Pengabdian Kepada Masyarakat UNSIQ*, vol. 9, no. 1, pp. 38–47, Dec. 2021.
- [6] M. Ekanto, "ANALISIS SENTIMEN ULASAN GOOGLE MAPS UNTUK REKOMENDASI TEMPAT WISATA DI YOGYAKARTA," *Jurnal Teknologi Informasi*, vol. 2, no. 2, pp. 55–116, Jul. 2024.
- [7] J. Ipawati, S. Saifulloh, and K. Kusnawi, "Analisis Sentimen Tempat Wisata Berdasarkan Ulasan pada Google Maps Menggunakan Algoritma Support Vector Machine," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 4, no. 1, pp. 247–256, Jan. 2024, doi: 10.57152/malcom.v4i1.1066.
- [8] I. Nabawi, Rudiman, and F. Yulianto, "3.+Ikhsan_ANALISIS+SENTIMEN+PADA+ULASAN+APLIKASI+GOOGLE+MAPS+TERHADAP+PELAYANAN+BADAN+PENYELenggara+JAMINAN+SOSIAL+(BPJS)+KESEHATAN+SAMARINDA+MENGGUNAKAN+METODE+K-NEAREST+NE," *Jurnal Keilmuan dan Aplikasi Bidang Teknik Informatika*, vol. 18, no. 2, pp. 100–116, Aug. 2024, Accessed: Jun. 09, 2025. [Online]. Available: <https://e-journal.upr.ac.id/index.php/JTI>
- [9] S. A. Rahmawati F, "Pemanfaatan Lokawisata Gedung Juang 45 sebagai Pemulihan Ekonomi Masyarakat Bekasi," *Jurnal Pariwisata Dan Budaya*, vol. 14, no. 1, pp. 76–87, Aug. 2023, doi: 10.31294/khi.v14i1.13039.
- [10] N. L. W. S. R. Ginantra, C. P. Yanti, G. D. Prasetya, I. B. G. Sarasvananda, and I. K. A. G. Wiguna, "Analisis Sentimen Ulasan Villa di Ubud Menggunakan Metode Naive Bayes, Decision Tree, dan K-NN," *Jurnal Nasional Pendidikan Teknik Informatika (JANAPATI)*, vol. 11, no. 3, pp. 205–215, Dec. 2022, doi: 10.23887/janapati.v11i3.49450.

- [11] N. L. W. S. R. Ginantra, C. P. Yanti, G. D. Prasetya, I. B. G. Sarasvananda, and I. K. A. G. Wiguna, "Analisis Sentimen Ulasan Villa di Ubud Menggunakan Metode Naive Bayes, Decision Tree, dan K-NN," *Jurnal Nasional Pendidikan Teknik Informatika (JANAPATI)*, vol. 11, no. 3, pp. 205–215, Dec. 2022, doi: 10.23887/janapati.v11i3.49450.
- [12] A. W. Sari, T. I. Hermanto, and M. Defriani, "Sentiment Analysis Of Tourist Reviews Using K-Nearest Neighbors Algorithm And Support Vector Machine," *Sinkron*, vol. 8, no. 3, pp. 1366–1378, Jul. 2023, doi: 10.33395/sinkron.v8i3.12447.
- [13] R. Oktafiani and R. Rianto, "Perbandingan Algoritma Support Vector Machine (SVM) dan Decision Tree untuk Sistem Rekomendasi Tempat Wisata," *Jurnal Nasional Teknologi dan Sistem Informasi*, vol. 9, no. 2, pp. 113–121, Aug. 2023, doi: 10.25077/teknosi.v9i2.2023.113-121.
- [14] Ahmad Bahar, Tri Astuti, and Primandani Arsi, "PERFORMANCE COMPARISON OF SVM, NAIVE BAYES, AND LOGISTIC REGRESSION CLASSIFICATION ALGORITHMS IN ANALYZING NOICE APP USER REVIEWS," *Jurnal Teknik Informatika (Jutif)*, vol. 5, no. 4, pp. 469–477, Aug. 2024, doi: 10.52436/1.jutif.2024.5.4.2061.
- [15] S. Sathyanarayanan, "Confusion Matrix-Based Performance Evaluation Metrics," *African Journal of Biomedical Research*, pp. 4023–4031, Nov. 2024, doi: 10.53555/AJBR.v27i4S.4345