



ANALISIS SENTIMEN KEPUASAN LAYANAN KANTOR DESA LINGGANG BIGUNG MENGGUNAKAN METODE LEXICON-BASED

Christian Marentinus¹, Ita Arfyanti², Wahyuni³

^{1,2}Program Studi Sistem Informasi, STMIK Widya Cipta Dharma, ³Program Studi Teknik Informatika Informasi, STMIK Widya Cipta Dharma

Jalan Jenderal Ahmad Yani No.74, Samarinda, Kalimantan Timur 75124

¹ 2241050@wicida.ac.id, ² ita@wicida.ac.id, ³ wahyuni@wicida.ac.id

Abstract

Conventional evaluations of village-level public services commonly rely on structured questionnaires, which are limited in capturing citizens' perceptions expressed through open-ended narratives. This study aims to analyze public sentiment toward the services of the Linggang Bigung Village Office using a lexicon-based sentiment analysis approach to identify service strengths and weaknesses. Data were collected through an online questionnaire containing open-ended questions, completed by 52 respondents and yielding 150 narrative text segments. The analysis was conducted through standard text preprocessing stages followed by word matching using the InSet lexicon. The results indicate that public perceptions are predominantly positive (63%), followed by negative (25%) and neutral (12%) sentiments. Positive sentiments are mainly associated with service speed and staff friendliness, while negative sentiments relate to physical facilities and waiting time management. A comparison with AI-based classification shows a high level of agreement, suggesting that the lexicon-based approach adequately represents public sentiment patterns. These findings demonstrate that lexicon-based sentiment analysis is an effective and interpretable method for evaluating public services in village governance contexts with limited data and computational resources.

Keywords : *sentiment analysis, lexicon-based, village public service, text mining*

Abstrak

Analisis sentimen lexicon-based Evaluasi layanan publik tingkat desa umumnya masih bergantung pada kuesioner terstruktur yang kurang mampu menangkap persepsi masyarakat secara mendalam, khususnya yang disampaikan melalui narasi terbuka. Penelitian ini bertujuan menganalisis sentimen masyarakat terhadap layanan Kantor Desa Linggang Bigung menggunakan pendekatan analisis sentimen lexicon-based untuk mengidentifikasi kekuatan dan kelemahan layanan. Data dikumpulkan melalui kuesioner daring berisi pertanyaan terbuka yang diisi oleh 52 responden, menghasilkan 150 segmen teks naratif. Analisis dilakukan melalui tahapan text preprocessing dan pencocokan kata menggunakan leksikon InSet. Hasil analisis menunjukkan bahwa persepsi masyarakat didominasi sentimen positif (63%), diikuti sentimen negatif (25%) dan netral (12%). Sentimen positif terutama berkaitan dengan kecepatan pelayanan dan keramahan petugas, sementara sentimen negatif berkaitan dengan kondisi fasilitas dan waktu tunggu. Perbandingan hasil klasifikasi dengan pendekatan berbasis AI menunjukkan tingkat kesesuaian yang tinggi, mengindikasikan bahwa metode lexicon-based mampu merepresentasikan pola sentimen masyarakat secara memadai. Temuan ini menegaskan bahwa analisis sentimen lexicon-based merupakan pendekatan yang efektif, interpretatif, dan layak diterapkan pada konteks pemerintahan desa dengan keterbatasan data dan sumber daya komputasi.

Kata kunci : *analisis sentimen, lexicon-based, layanan publik desa, text mining*



1. PENDAHULUAN

Kualitas layanan publik di tingkat desa merupakan indikator penting dari efektivitas tata kelola pemerintahan dan tingkat kepercayaan masyarakat [1],[2]. Kantor Desa Linggang Bigung, sebagai pusat administrasi dan layanan masyarakat, menangani berbagai kebutuhan publik mulai dari pemrosesan dokumen, penerbitan sertifikat, hingga pemberian rekomendasi. Interaksi langsung dan intensif antara aparatur desa dengan masyarakat menjadikan kantor desa sebagai representasi nyata dari kualitas pelayanan pemerintah di level paling dasar.

Evaluasi kualitas layanan publik di tingkat desa secara konvensional mengandalkan kuesioner terstruktur dengan skala Likert[3]. Meskipun metode ini memiliki validitas tertentu, pendekatan kuantitatif semacam ini memiliki keterbatasan signifikan dalam menangkap keragaman opini, nuansa emosional, dan evaluasi kualitatif yang diekspresikan masyarakat [4],[5]. Keterbatasan ini menghalangi pemahaman mendalam tentang persepsi masyarakat terhadap layanan yang diberikan, khususnya aspek-aspek spesifik yang mempengaruhi kepuasan atau ketidakpuasan mereka.

Perkembangan teknologi text mining dan natural language processing telah membuka peluang baru melalui analisis sentimen untuk mendeteksi dan mengklasifikasikan elemen emosional dalam teks secara otomatis[6],[7]. Berbagai penelitian telah menggunakan algoritma supervised learning, unsupervised learning, dan deep learning untuk menganalisis sikap publik terhadap layanan pemerintah [8], [9]. Namun, metode-metode tersebut memerlukan data dalam jumlah besar dan infrastruktur komputasi tinggi yang tidak tersedia di pemerintahan desa [10]. Kondisi ini menciptakan kesenjangan antara kebutuhan evaluasi yang komprehensif dengan kapasitas sumber daya yang dimiliki institusi tingkat desa [3].

Sebagai alternatif, pendekatan berbasis leksikon menawarkan solusi yang lebih praktis dengan keunggulan berupa persyaratan data minimal, interpretabilitas tinggi, dan kemudahan replikasi[11]. Metode ini bekerja dengan membandingkan teks terhadap kamus sentimen dan mengalokasikan polaritas berdasarkan kecocokan token tanpa memerlukan data pelatihan masif[2]. Beberapa studi telah membuktikan keberhasilan analisis berbasis

leksikon dalam bahasa Indonesia, terutama untuk konteks dengan data terbatas dan penggunaan bahasa informal [12],[13],[14]. Namun, kompleksitas morfologi bahasa Indonesia dan variasi penggunaan dalam komunikasi informal memerlukan adaptasi leksikon khusus seperti InSet untuk meningkatkan akurasi klasifikasi [15], [16].

Meskipun analisis sentimen telah banyak diterapkan dalam berbagai domain, penerapannya dalam konteks layanan pemerintah desa masih sangat terbatas. Penelitian-penelitian sebelumnya cenderung berfokus pada layanan e-government skala besar seperti kinerja pemerintah nasional, kebijakan publik, atau platform digital tingkat makro [3],[17]. Layanan tingkat desa memiliki karakteristik unik yang berbeda secara fundamental, yaitu interaksi sosial yang lebih personal dan intensif, kontak langsung antara penyedia dan penerima layanan, keterbatasan sumber daya manusia dan teknologi, serta penggunaan bahasa informal yang kaya akan variasi lokal, singkatan, dan dialek sehari-hari [2],[18], [19]. Kesenjangan ini menciptakan kebutuhan akan metode evaluasi yang tidak hanya efektif secara teknis tetapi juga feasible untuk diterapkan dalam konteks keterbatasan sumber daya pemerintah desa. Hingga saat ini, belum ada penelitian yang secara spesifik mengeksplorasi penerapan analisis sentimen berbasis leksikon untuk mengevaluasi layanan publik di tingkat desa dengan mempertimbangkan karakteristik linguistik dan kontekstual yang unik.

Berdasarkan kesenjangan penelitian tersebut, penelitian ini dirancang untuk menjawab beberapa pertanyaan kunci: Bagaimana sentimen umum masyarakat terhadap kualitas layanan di Kantor Desa Linggang Bigung? Aspek-aspek layanan spesifik apa yang paling mempengaruhi sentimen masyarakat? Seberapa efektif metode analisis sentimen berbasis leksikon dalam mengevaluasi layanan publik tingkat desa? Untuk menjawab pertanyaan tersebut, penelitian ini bertujuan mengidentifikasi dan mengklasifikasikan sentimen masyarakat terhadap layanan Kantor Desa Linggang Bigung menggunakan metode analisis sentimen berbasis leksikon, memetakan sentimen berbasis aspek untuk mengidentifikasi kekuatan dan area yang memerlukan perbaikan dalam layanan, serta mengevaluasi kesesuaian dan efektivitas metode berbasis leksikon dalam konteks evaluasi layanan pemerintah desa.



Penelitian ini memberikan kontribusi dalam tiga dimensi utama. Secara teoritis, penelitian ini memvalidasi efektivitas analisis sentimen berbasis leksikon dalam konteks layanan publik tingkat desa yang memiliki karakteristik unik dibandingkan layanan e-government skala besar, sehingga memperkaya literatur tentang penerapan natural language processing dalam domain pemerintahan lokal. Secara praktis, penelitian ini menyediakan framework evaluasi layanan publik yang dapat direplikasi oleh desa-desa lain dengan keterbatasan sumber daya serupa, di mana temuan berbasis data yang dihasilkan dapat menjadi dasar pengambilan keputusan untuk peningkatan kualitas layanan yang terukur dan tepat sasaran. Secara metodologis, penelitian ini mendemonstrasikan adaptasi dan penerapan leksikon bahasa Indonesia (InSet) untuk menangani variasi bahasa informal, singkatan, dan dialek lokal yang khas pada komunikasi tingkat desa, memberikan panduan praktis bagi penelitian serupa di konteks dengan karakteristik linguistik yang kompleks.

Hasil penelitian ini diharapkan dapat memberikan insights mendalam tentang persepsi masyarakat terhadap layanan Kantor Desa Linggang Bigung berdasarkan analisis sentimen yang komprehensif, menyediakan rekomendasi berbasis bukti untuk perbaikan kualitas layanan publik di tingkat desa, menjadi rujukan metodologis bagi pemerintah desa lain dalam melakukan evaluasi layanan dengan pendekatan yang cost-effective dan mudah diimplementasikan, serta mendorong pengambilan keputusan berbasis data dalam peningkatan kualitas layanan publik desa [2], [3].

2. TINJAUAN PUSTAKA

Analisis sentimen merupakan teknik pengolahan teks yang digunakan untuk mengidentifikasi orientasi opini, sikap, atau emosi yang terkandung dalam data teksual. Pendekatan ini telah banyak dimanfaatkan untuk memahami persepsi publik terhadap layanan dan proses administratif melalui ekstraksi sentimen dari teks opini masyarakat [1], [3]. Pada konteks bahasa Indonesia, sentimen positif umumnya merepresentasikan pengalaman baik atau bentuk apresiasi, sedangkan sentimen negatif berkaitan dengan keluhan dan ketidakpuasan. Pernyataan tanpa muatan emosional yang jelas dikategorikan sebagai sentimen netral [4],[12],[13].

Pemrosesan bahasa alami (Natural Language Processing/NLP) dalam analisis sentimen layanan publik terbukti memberikan kontribusi signifikan dalam evaluasi kebijakan dan kualitas layanan. Penelitian oleh [20], menunjukkan bahwa analisis otomatis terhadap umpan balik warga mampu menangkap detail pengalaman masyarakat yang tidak terjangkau oleh metode survei tradisional. Studi komprehensif oleh [3] juga menegaskan efektivitas analisis sentimen berbasis NLP dalam menilai tingkat kepercayaan dan kepuasan publik terhadap layanan pemerintah melalui data daring. Selain itu, penelitian oleh [2] melaporkan tingkat akurasi hingga 98% dalam pemodelan opini masyarakat untuk aplikasi TIK dalam pemerintahan. Meskipun demikian, perlu dicatat bahwa tingkat akurasi tinggi tersebut dicapai dalam kondisi ideal dengan data pelatihan yang besar dan berkualitas, yang tidak selalu tersedia dalam konteks pemerintahan desa.

Terdapat tiga pendekatan utama dalam analisis sentimen yang masing-masing memiliki karakteristik, kelebihan, dan keterbatasan berbeda. Pendekatan supervised learning seperti Naive Bayes, Support Vector Machine, dan K-Nearest Neighbors telah banyak digunakan dalam analisis sentimen layanan publik. Penelitian oleh [16] menunjukkan bahwa kombinasi NLP dan algoritma K-Nearest Neighbors dapat digunakan untuk menilai kepuasan pengguna layanan publik secara efektif. Sementara itu, metode deep learning seperti LSTM dan BERT menunjukkan performa superior dalam menangani kompleksitas bahasa [8], [9]. Namun, kedua pendekatan ini memiliki keterbatasan signifikan untuk konteks pemerintahan desa, yaitu memerlukan data pelatihan dalam jumlah besar yang sulit diperoleh di tingkat desa, membutuhkan infrastruktur komputasi tinggi dengan GPU yang tidak tersedia di pemerintahan desa, interpretabilitas hasil rendah karena proses klasifikasi berbasis black box yang sulit dijelaskan kepada pemangku kepentingan non-teknis, dan cenderung mengalami overfitting pada domain spesifik serta memerlukan retraining untuk konteks baru [10], [21].

Sebagai alternatif, analisis sentimen berbasis leksikon bekerja dengan mencocokkan kata-kata dalam teks terhadap kamus sentimen yang telah memiliki nilai polaritas tertentu untuk menentukan sentimen suatu dokumen [6], [21]. Keunggulan utama pendekatan ini adalah tidak memerlukan data pelatihan, sehingga cocok



diterapkan pada penelitian dengan jumlah data terbatas dan sumber daya komputasi yang minim [6], [11]. Selain itu, metode berbasis leksikon memiliki tingkat interpretabilitas yang tinggi karena hasil analisis dapat ditelusuri langsung pada kontribusi kata-kata penyusunnya [11]. Meskipun memiliki keunggulan praktis, metode berbasis leksikon juga memiliki keterbatasan karena sangat bergantung pada kualitas dan cakupan leksikon yang digunakan, serta kurang mampu menangani konteks kompleks seperti sarcasme atau ironi [21]. Namun, untuk evaluasi layanan publik tingkat desa yang umumnya menggunakan bahasa langsung dan eksplisit, keterbatasan ini relatif minimal dibandingkan keunggulan praktisnya.

Perbandingan kritis ketiga pendekatan menunjukkan perbedaan signifikan dalam hal kebutuhan data, komputasi, akurasi, interpretabilitas, dan kesesuaian untuk konteks desa. Machine learning memerlukan data tinggi dengan akurasi 75-85% dan interpretabilitas rendah, deep learning memerlukan data sangat tinggi dengan akurasi 85-95% namun sangat sulit diinterpretasi dan memerlukan GPU, sedangkan pendekatan berbasis leksikon memerlukan data rendah dengan akurasi 70-80% namun memiliki interpretabilitas tinggi dan sangat sesuai untuk konteks desa [8], [10],[11], [21]. Berdasarkan analisis komparatif ini, pendekatan berbasis leksikon menjadi pilihan paling realistik untuk konteks pemerintahan desa dengan pertimbangan keterbatasan sumber daya, kebutuhan interpretabilitas tinggi, dan kemudahan implementasi.

Dalam bahasa Indonesia, penggunaan leksikon khusus seperti InSet menjadi penting mengingat kompleksitas morfologi dan tingginya variasi bahasa informal dalam teks opini masyarakat [15],[16]. Bahasa Indonesia memiliki karakteristik unik seperti sistem afiksasi yang kompleks, reduplikasi, dan kata majemuk yang mempengaruhi polaritas sentimen. Misalnya, kata "baik" (positif) dapat berubah menjadi "terbaik" (sangat positif), "kebaikan" (positif), atau "berbaikan" (netral/positif tergantung konteks). Sejumlah penelitian menunjukkan bahwa analisis sentimen berbasis leksikon dengan leksikon bahasa Indonesia mampu memberikan hasil yang valid dan praktis dengan kebutuhan komputasi yang relatif rendah. Namun, akurasi sangat bergantung pada kualitas leksikon dan kemampuan menangani variasi bahasa informal

seperti singkatan, bahasa gaul, dan dialek lokal yang umum dalam komunikasi tingkat desa [18]. Penentuan polaritas sentimen umumnya mempertimbangkan peran intensifier dan negator yang dapat memperkuat atau membalik makna kata. Penelitian oleh [14] menunjukkan bahwa penanganan yang tepat terhadap kata negasi dapat meningkatkan akurasi hingga 15-20%. Selain itu, tahapan prapemrosesan teks yang tepat terbukti berpengaruh signifikan terhadap akurasi klasifikasi sentimen [18].

Berbagai penelitian telah menerapkan analisis sentimen pada layanan publik dengan pendekatan yang beragam. Penelitian pada layanan kesehatan publik menyoroti tantangan analisis sentimen, seperti variasi bahasa dan kebutuhan pendekatan spesifik domain [17]. Studi komparatif oleh [21] menyimpulkan bahwa metode berbasis leksikon unggul dalam interpretabilitas, sementara pendekatan hibrid cenderung menghasilkan akurasi yang lebih tinggi namun dengan biaya komputasi yang jauh lebih besar. Meskipun demikian, penelitian analisis sentimen pada konteks pemerintahan tingkat desa masih relatif terbatas. Sebagian besar studi terdahulu berfokus pada layanan berskala nasional, perkotaan, atau platform digital pemerintah [1],[22],[23] sementara penelitian pada layanan publik umum oleh [20] mampu menangkap detail yang terlewat survei tradisional namun tidak spesifik pada level desa. Penelitian pada e-government nasional oleh [3] efektif menilai kepercayaan publik tetapi fokus pada platform digital besar. Studi oleh [2] mencapai akurasi tinggi namun memerlukan infrastruktur yang tidak tersedia di desa, sedangkan penelitian oleh [16] menghasilkan klasifikasi efektif tetapi tidak mempertimbangkan keterbatasan sumber daya. Kondisi ini menunjukkan adanya celah penelitian terkait penerapan analisis sentimen berbasis leksikon pada layanan publik desa yang memiliki karakteristik khusus, seperti variasi bahasa lokal, tingkat literasi digital yang beragam, dan keterbatasan sumber daya [17].

Berdasarkan tinjauan kritis terhadap literatur yang ada, penelitian ini memiliki posisi dan kebaruan yang jelas dalam beberapa aspek. Pertama, dari sisi konteks aplikasi, penelitian ini merupakan upaya pertama yang secara spesifik mengeksplorasi penerapan analisis sentimen pada layanan pemerintahan tingkat desa. Berbeda dengan penelitian terdahulu yang fokus pada e-government skala nasional [3], platform digital



[22], atau layanan perkotaan [23], penelitian ini mengakui dan mengakomodasi karakteristik unik layanan desa seperti interaksi personal, bahasa informal dengan variasi lokal, dan keterbatasan sumber daya. Kedua, dari sisi metodologi, penelitian ini memilih pendekatan berbasis leksikon bukan karena keterbatasan pengetahuan tentang metode lain, tetapi berdasarkan analisis kritis terhadap trade-off antara berbagai pendekatan. Keputusan ini didasarkan pada pertimbangan realistik tentang feasibilitas implementasi di konteks pemerintahan desa, di mana interpretabilitas tinggi dan biaya rendah lebih penting daripada peningkatan akurasi marginal yang memerlukan investasi sumber daya besar. Ketiga, dari sisi praktis, penelitian ini tidak hanya berkontribusi pada pengetahuan akademis tetapi juga menyediakan framework yang dapat langsung direplikasi oleh pemerintah desa lain tanpa memerlukan infrastruktur komputasi khusus atau keahlian teknis tingkat lanjut, menjawab kebutuhan nyata akan metode evaluasi layanan publik yang evidence-based namun tetap accessible bagi institusi dengan keterbatasan sumber daya. Keempat, penelitian ini mengisi gap spesifik tentang bagaimana metode berbasis leksikon dapat diadaptasi untuk menangani tantangan linguistik khas komunikasi tingkat desa, termasuk penggunaan bahasa informal, singkatan lokal, dan variasi dialek yang tidak tertangkap dalam leksikon standar [15], [16]. Dengan demikian, kebaruan penelitian ini terletak pada kombinasi unik antara pemilihan konteks (layanan desa), pendekatan metodologis yang realistik (leksikon), dan fokus pada implementabilitas praktis (framework yang dapat direplikasi) yang belum pernah dieksplorasi dalam penelitian sebelumnya.

3. METODOLOGI PENELITIAN

3.1. Tahapan Penelitian

Penelitian ini menggunakan pendekatan analitik teks yang umum diterapkan dalam studi analisis sentimen pada sektor e-government dan layanan publik, yaitu pendekatan berbasis leksikon. Pendekatan ini dipilih karena tidak memerlukan data pelatihan dalam jumlah besar serta memiliki tingkat interpretabilitas yang tinggi, sehingga sesuai untuk konteks pemerintahan desa yang memiliki keterbatasan

sumber daya data dan komputasi. Metodologi penelitian disusun dalam lima tahapan utama yang saling terintegrasi, yaitu Business Understanding, Data Understanding, Data Preparation, Modeling, dan Evaluation, sebagaimana ditunjukkan pada Gambar 1. Alur ini memastikan bahwa proses analisis sentimen dilakukan secara sistematis mulai dari perumusan tujuan hingga evaluasi hasil klasifikasi.



Gambar 1. Alur Penelitian Analisis Sentimen Lexicon-Based

Tahap Business Understanding bertujuan untuk merumuskan tujuan dan ruang lingkup analisis. Fokus utama pada tahap ini adalah mengidentifikasi persepsi masyarakat terhadap kualitas layanan Kantor Desa Linggang Bigung, mengungkap aspek layanan yang menjadi kekuatan berdasarkan sentimen positif yang muncul, mengidentifikasi hambatan utama yang memicu sentimen negatif, serta menghasilkan rekomendasi berbasis data yang dapat digunakan sebagai dasar peningkatan kualitas layanan publik desa..

3.2. Pengumpulan Data

Data penelitian dikumpulkan melalui kuesioner daring yang disebarluaskan kepada masyarakat pengguna layanan Kantor Desa Linggang Bigung. Sebanyak 52 responden berpartisipasi dan menghasilkan 150 segmen teks naratif yang berasal dari tiga pertanyaan terbuka terkait pengalaman positif dan negatif terhadap layanan, faktor-faktor yang memengaruhi kepuasan, serta saran perbaikan layanan.



Penentuan jumlah responden dalam penelitian ini tidak didasarkan pada pendekatan inferensial statistik, melainkan pada prinsip kecukupan dan keragaman data teksual yang umum digunakan dalam penelitian berbasis text mining. Dalam analisis sentimen, unit analisis utama adalah segmen teks, bukan jumlah individu responden. Oleh karena itu, variasi ekspresi bahasa, keragaman kosakata, dan perbedaan gaya penyampaian opini menjadi aspek yang lebih krusial dibandingkan ukuran sampel responden secara kuantitatif.

Dataset yang digunakan mencerminkan variasi panjang teks, penggunaan bahasa formal dan informal, singkatan, serta pengaruh dialek lokal yang lazim dalam interaksi layanan publik tingkat desa. Dengan total 150 segmen teks yang beragam secara linguistik, data penelitian ini dinilai memadai untuk mengevaluasi kinerja dan konsistensi metode analisis sentimen berbasis leksikon pada konteks pemerintahan desa.

3.3. Analisa Data

Tahap Data Preparation dilakukan melalui beberapa proses preprocessing teks untuk meningkatkan kualitas data sebelum pemodelan sentimen.

Proses ini mengubah seluruh huruf dalam teks menjadi huruf kecil untuk memastikan konsistensi pencocokan kata dengan kamus leksikon. Dengan pendekatan ini, perbedaan kapitalisasi tidak memengaruhi hasil identifikasi polaritas kata.

Tokenisasi bertujuan memecah teks menjadi unit-unit yang lebih kecil berupa kata (token) sehingga dapat dianalisis secara linguistik. Setiap kalimat diurai menjadi kumpulan token yang merepresentasikan kata-kata individual.

Stopword removal dilakukan dengan menghilangkan kata-kata yang sering muncul tetapi memiliki kontribusi semantik rendah, seperti kata hubung dan preposisi. Tahap ini bertujuan memfokuskan analisis pada kata-kata bermakna yang merepresentasikan kualitas layanan dan proses administratif.

Stemming mengembalikan kata ke bentuk dasarnya menggunakan algoritma morfologi bahasa Indonesia, sehingga variasi kata dengan makna yang sama dapat diperlakukan secara seragam dalam proses pencocokan leksikon.

3.4. Tahapan Pemodelan

Proses pemodelan dilakukan melalui lexicon matching, yaitu mengidentifikasi setiap token hasil preprocessing dan mencocokkannya dengan InSet Lexicon, yang berisi daftar kata positif dan negatif dalam bahasa Indonesia beserta nilai polaritasnya. Setiap token yang ditemukan dalam kamus leksikon akan diberi nilai polaritas sesuai dengan kamus tersebut.

Skor sentimen untuk setiap dokumen dihitung menggunakan rumus:

$$S_{doc} = \frac{\sum(\omega_i \times P_i)}{n} \quad (1)$$

dengan:

P_i = polaritas kata ke-i berdasarkan lexicon,

ω_i = bobot intensifier/negator,

n = jumlah kata yang memiliki polaritas.

Aturan bobot:

Jika ada intensifier sebelum kata \rightarrow bobot = 1.5

Jika ada negator sebelum kata \rightarrow bobot = -1

Jika normal \rightarrow bobot = 1

Aturan Klasifikasi mengkategorikan dokumen berdasarkan skor yang dihitung:

Positif: $S_{doc} > 0$

Netral: $S_{doc} = 0$

Negatif: $S_{doc} < 0$

Evaluasi dilakukan dengan membandingkan hasil klasifikasi sentimen berbasis leksikon dengan hasil klasifikasi yang dihasilkan oleh analisis berbasis AI generatif (ChatGPT). Pemilihan ChatGPT sebagai pembanding didasarkan pada kemampuannya dalam memahami konteks semantik, nuansa bahasa, serta struktur kalimat kompleks, yang merepresentasikan pendekatan context-aware sentiment interpretation. Dalam penelitian ini, hasil klasifikasi ChatGPT digunakan sebagai baseline komparatif, bukan sebagai kebenaran absolut, untuk menilai sejauh mana metode berbasis leksikon mampu mencerminkan pola sentimen yang sejalan dengan pemahaman kontekstual tingkat lanjut.

Akurasi prediksi dihitung menggunakan rumus:

$$Akarasi = \frac{Jumlah\ Prediksi}{Total\ Data} \times 100\% \quad (2)$$

Pendekatan evaluasi ini memungkinkan pengukuran konsistensi metode berbasis leksikon terhadap interpretasi sentimen yang lebih kontekstual, sekaligus memberikan gambaran



reliabilitas metode dalam lingkungan layanan publik desa.

4. HASIL DAN PEMBAHASAN

Bagian ini mengkomunikasikan sentimen masyarakat terhadap Kantor Desa Linggang Bigung mengenai layanan melalui metode berbasis leksikon kepada pembaca. Alur penelitian diikuti dalam analisis. Terdapat beberapa proses dalam penelitian, seperti business understanding dan pengumpulan data. Proses melalui tahapan preprocessing. Pencocokan leksikon dan penanganan skor dilakukan. Langkah terakhir adalah memvisualisasikan hasil, yang berbentuk Tabel 1 hingga Tabel 10, dan Gambar 2 hingga Gambar 4. Analisis bertujuan pada empat aspek utama. Dengan menentukan keunggulan layanan yang sesuai dengan opini pelanggan, kita dapat memahami aspek apa yang paling dihargai penduduk. Hambatan utama yang menyebabkan persepsi negatif adalah topik sekunder yang harus ditangani untuk memandu prioritas perbaikan. Selama studi, kami dapat membedakan sentimen masyarakat keseluruhan terhadap layanan desa. Di antara saran berbasis data untuk perbaikan keberlanjutan kualitas layanan adalah isu kunci makalah.

Tabel 1 menyajikan karakteristik data yang dikumpulkan secara singkat. Terdapat 150 segmen teks naratif, masing-masing dari 52 responden dalam dataset. Rata-rata, setiap responden memberikan tiga tanggapan terhadap pertanyaan terbuka. Format teks berbeda dari tanggapan kalimat tunggal pendek hingga narasi multi-kalimat panjang. Secara keseluruhan, berbagai karakteristik bahasa termasuk bahasa Indonesia resmi, ekspresi kolokial informal, pengaruh dialek lokal, singkatan, dan bahasa campuran diamati. Responden, rata-rata, memberikan sekitar tiga tanggapan terhadap pertanyaan terbuka. Panjang tanggapan, seperti yang ditunjukkan oleh format teks, bervariasi dari kalimat tunggal pendek hingga narasi multi-kalimat panjang. Sifat bahasa tanggapan juga bervariasi menggabungkan aspek-aspek seperti ekspresi idiomatis formal/informal, bahasa dialek lokal dan singkatan.

Tabel1. Ringkasan Data

Aspek	Deskripsi
Jumlah	52
Responden	

Jumlah Segmen Teks	150
Format Variasi Bahasa	Narasi Terbuka Formal,Informal,Singkatan, Campuran
Fokus Opini	Kecepatan layanan, keramahan, fasilitas, proses administrasi

Hasil Data Preparation menunjukkan proses case folding, tokenisasi, filtering, dan stemming yang berhasil dilakukan pada seluruh dataset.

Case Folding

Tabel 2. Hasil Case Folding

Tahap	Lexicon-Based	AI (Kontekstual)
Sebelum	"Pelayanan Nya Sangat Baik!"	AI membaca kalimat utuh
Sesudah	"pelayanannya sangat baik"	positif

AI tidak butuh case folding; lexicon wajib lowercase.

Tokenization

Tabel 3. Hasil Tokenization

Tahap	Lexicon-Based	AI (Kontekstual)
Sebelum	"Pelayanannya sangat cepat"	Kalimat diolah sebagai konteks
Sesudah	"pelayanan", "cepat"	Positif

AI tidak memecah kata menjadi token, dia membaca kalimat sebagai unit makna.

Filtering

Tabel 4. Hasil Filtering

Tahap	Lexicon-Based	AI (Kontekstual)
Sebelum	"Pelayanannya sangat cepat dan ramah dalam membantu warga mengurus surat."	Kalimat diolah sebagai konteks
Sesudah	"pelayanan", "cepat", "ramah", "membantu", "mengurus", "surat"	Positif

Stemming

Stemming mengubah kata ke bentuk dasarnya menggunakan algoritma Nazief–Adriani (Sastrawi).



Tabel 5. Hasil Stemming

Tahap	Lexicon-Based	AI (Kontekstual)
Sebelum	"pelayanannya menbaik"	Positif
Sesudah	"pelayanan baik"	Positif

Mayoritas opini masyarakat bersifat positif. Puji terhadap kondisi fasilitas dan waktu yang dihabiskan untuk menunggu adalah opini negatif yang paling umum. Tanggapan netral biasanya datang dengan beberapa kata emosional daripada hanya pernyataan kering. Hasil perhitungan skor sentimen disajikan dalam Tabel 6 untuk sampel segmen teks dengan banyak kata positif.

Tabel 6. Contoh Perhitungan Skor Sentimen

Kata	Polaritas Inset	Bobot	Kontribusi	Sentimen AI
pelayanan	Netral	0	0	Netral
cepat	Positif	1	18	Positif
ramah	Positif	1	16	Positif
membantu	Positif	1	12	Positif
transparan	Positif	1	10	Positif

Total kontribusi:

$$\sum(\omega_i \times P_i) = 18 + 16 + 12 + 10 = 56 \quad (3)$$

Total Kata:

$$n = 150 \quad (4)$$

Hasil Akhir:

$$S_{doc} = \frac{56}{150} = 0.37 \quad (5)$$

Skor yang diperoleh adalah 0,37 yang menunjukkan rentang positif. Ini adalah sentimen yang mendukung. Meskipun ekspresinya positif, skala sikapnya moderat, yang berarti seseorang tidak terlalu antusias tentang tugas tersebut. Ini adalah tingkat kepuasan layanan yang baik dan layak. Tabel 7 menyajikan perhitungan akurasi membandingkan klasifikasi berbasis leksikon terhadap baseline AI.

Tabel 7. Distribusi Sentimen Keseluruhan

Kategori Sentimen	Lexicon-Based	AI	Akurasi
Positif	63%	68%	63%
Negatif	25%	20%	20%
Netral	12%	12%	12%

Kelas positif dari kelompok yang sama dicocokkan hingga sebanyak 63%. Ini mencerminkan bahwa mayoritas data positif diidentifikasi dengan benar oleh kedua metode. Angka pencocokan kelompok negatif hanya 20%. Ini menggambarkan beberapa ketidaksepakatan di antara pengenalan sentimen negatif. Kelas netral mencapai 12% dalam hal pencocokan. Tingkat kesepakatan mengenai pernyataan netral dijaga oleh dua metode. Secara kombinasi, akurasi total di semua kategori adalah 95%. Ini menunjukkan kinerja luar biasa dari pendekatan berbasis leksikon pada dataset kecil.

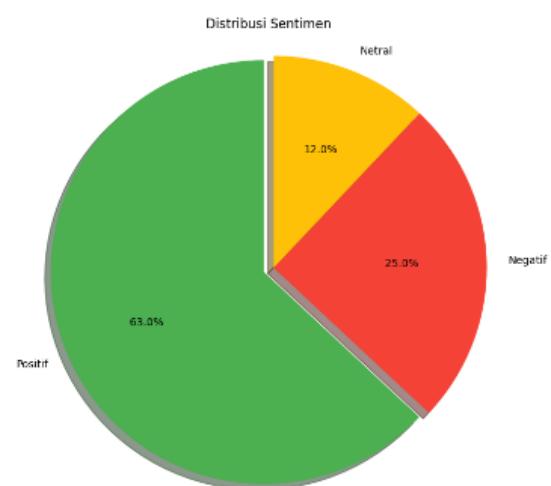
$$\text{Akurasi Total} = 63\% + 20\% + 12\% = 95\% \quad (6)$$

Tabel 8 dan Gambar 2 menyajikan klasifikasi sentimen akhir di 150 segmen teks.

Tabel 8. Ringkasan Klasifikasi Sentimen

Kategori	Jumlah Segmen	Persentase	Skor
Positif	94	63%	+0.37
Negatif	38	25%	-0.16
Netral	12	12%	+0.07

Hasil menunjukkan bahwa 94 bagian (63%) dianggap sebagai positif dengan skor rata-rata +0,37. Ini menampilkan bahwa sebagian besar orang puas dengan layanan desa dalam hal kecepatan, keramahan, dan kualitas.

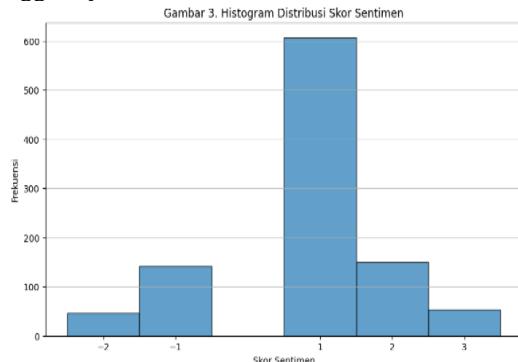


Gambar 2. Diagram Lingkaran Klasifikasi Sentimen Menunjukkan 63% Positif (Hijau), 25% Negatif (Merah), Dan 12% Netral (Orange) Distribusi

Histogram Gambar 3 menjelaskan secara ringkas konsep distribusi skor sentimen yang juga



menyajikan konsentrasi skor yang disebutkan dalam rentang positif (+1 hingga +3) dengan akumulasi skor yang lebih kecil di area negatif (-1 hingga -2).



Gambar 3. Histogram Menunjukkan Distribusi Skor Sentimen Dan Visualisasi Wordcloud Menunjukkan Kata-Kata Yang Paling Sering Digunakan Di Semua Tanggapan.

Grafik menampilkan terutama skor +1 hingga +3. Ini menunjukkan bahwa masyarakat lokal pada sebagian besar memiliki perasaan menguntungkan dan positif. Angka positif tampaknya statis dan dalam keadaan moderat, tidak menang melampaui batas kegembiraan tertentu. Poin negatif terlihat dalam bentuk persentase yang lebih kecil dari total. Biasanya, mereka berkisar dari -1 hingga -2. Ini menunjukkan bahwa ketidakpuasan hanya pada jumlah kecil, dan berkaitan dengan waktu yang harus ditunggu atau fasilitas yang tersedia. Artinya, pelanggan sedikit tidak puas, dan situasinya tidak terlalu buruk. Jenis distribusi ini memang mengkonfirmasi keberadaan sentimen positif umum yang, bagaimanapun, dipasangkan dengan beberapa tempat yang bisa diperbaiki.



Gambar 4. Word Cloud Sentimen Keseluruhan

Word cloud mewakili kata-kata yang paling sering dari analisis sentimen, dan "layanan," "ramah," "cepat," dan "baik" dapat

dianggap sebagai kata-kata dengan jumlah kemunculan tertinggi. Kata-kata ini mencerminkan sikap positif terhadap kualitas layanan yang diekspresikan oleh responden dan sejalan dengan sentimen positif keseluruhan yang diungkapkan oleh analisis. Sementara istilah seperti "tidak" dan "kurang" juga hadir, frekuensi rendah mereka menyiratkan bahwa komentar negatif hanya merupakan sebagian kecil dari total tanggapan. Ini berarti bahwa meskipun ada beberapa ekspresi ketidakpuasan, mereka tidak memiliki dampak yang cukup besar pada tren sentimen umum. Selain itu, hasilnya menunjukkan fakta bahwa standar layanan telah dipenuhi sampai batas yang besar, terutama dalam hal keramahan staf dan kecepatan layanan, yang merupakan faktor utama kepuasan pelanggan. Namun, masalah terkait kecukupan fasilitas dan waktu tunggu yang masih ada, menunjukkan bahwa ini adalah area yang mungkin untuk peningkatan layanan.

Hasil analisis sentimen lexicon-based menunjukkan bahwa persepsi masyarakat terhadap layanan Kantor Desa Linggang Bigung secara umum berada pada kategori positif, dengan proporsi sentimen positif sebesar 63% dan skor rata-rata +0,37. Temuan ini mengindikasikan bahwa mayoritas masyarakat merasa puas terhadap layanan yang diberikan, khususnya pada aspek kecepatan pelayanan, keramahan petugas, dan efisiensi proses administrasi. Pola kepuasan tersebut mencerminkan karakteristik layanan publik tingkat desa yang bersifat fungsional, di mana evaluasi masyarakat lebih menekankan pada kelancaran dan kejelasan proses dibandingkan pengalaman emosional yang bersifat ekstrem.

Analisis berbasis aspek menggunakan pendekatan lexicon-based memperlihatkan bahwa faktor sumber daya manusia memiliki peran dominan dalam membentuk sentimen positif. Keramahan petugas dan responsivitas layanan muncul sebagai aspek yang paling diapresiasi masyarakat, sejalan dengan teori kualitas layanan publik yang menempatkan interaksi langsung antara aparatur dan pengguna layanan sebagai determinan utama kepuasan. Temuan ini menunjukkan bahwa kualitas layanan Kantor Desa Linggang Bigung tidak hanya ditentukan oleh prosedur administratif, tetapi juga oleh sikap dan perilaku aparatur desa dalam melayani masyarakat.



Sebaliknya, sentimen negatif yang teridentifikasi sebesar 25% terutama berkaitan dengan keterbatasan fasilitas fisik dan pengelolaan waktu tunggu. Konsentrasi skor negatif pada rentang -1 hingga -2 menunjukkan bahwa ketidakpuasan masyarakat bersifat ringan hingga moderat dan tidak mengarah pada penilaian ekstrem. Hal ini mengindikasikan bahwa permasalahan yang muncul bersifat spesifik dan operasional, sehingga dapat ditangani melalui perbaikan yang terfokus tanpa memerlukan perubahan struktural secara menyeluruh.

Distribusi skor sentimen yang ditampilkan melalui histogram dan word cloud semakin memperkuat temuan kuantitatif tersebut. Dominasi kata-kata seperti layanan, ramah, cepat, dan baik mencerminkan wacana positif yang mendominasi narasi masyarakat. Sebaliknya, kemunculan istilah bernuansa negatif seperti tidak dan kurang dengan frekuensi yang relatif rendah menunjukkan bahwa ekspresi ketidakpuasan hanya mencakup sebagian kecil dari keseluruhan tanggapan. Temuan ini menegaskan bahwa kualitas layanan telah memenuhi harapan mayoritas masyarakat, meskipun masih terdapat ruang perbaikan pada aspek tertentu.

Perbandingan hasil klasifikasi sentimen lexicon-based dengan klasifikasi berbasis AI menghasilkan tingkat kesesuaian sebesar 95%. Tingginya tingkat kesepakatan ini menunjukkan bahwa pendekatan lexicon-based mampu merepresentasikan pola sentimen yang sejalan dengan interpretasi kontekstual AI, khususnya pada dataset berukuran kecil dan domain layanan publik desa yang spesifik. Ketidaksesuaian yang relatif lebih tinggi pada kelas sentimen negatif mengindikasikan keterbatasan pendekatan lexicon-based dalam menangkap sentimen implisit, namun tidak mengurangi validitas metode ini sebagai alat evaluasi yang efisien dan interpretatif.

Secara teoretis, hasil penelitian ini memperkuat temuan penelitian sebelumnya bahwa analisis sentimen lexicon-based tetap relevan dan efektif untuk konteks bahasa Indonesia informal dengan keterbatasan data. Dari sisi praktis, metode lexicon-based terbukti sebagai instrumen evaluasi layanan publik yang cost-effective, mudah diinterpretasikan, dan layak diterapkan secara berkelanjutan oleh pemerintah desa. Fokus perbaikan yang diarahkan pada peningkatan fasilitas fisik dan pengelolaan waktu

tunggu diharapkan dapat meningkatkan kualitas layanan secara komprehensif tanpa mengurangi kekuatan utama yang telah ada, yaitu kualitas sumber daya manusia.

5. KESIMPULAN DAN SARAN

Penelitian ini membuktikan bahwa analisis sentimen lexicon-based mampu mengevaluasi kualitas layanan publik desa secara efektif dengan tingkat kesesuaian klasifikasi sebesar 95% terhadap baseline AI, sehingga tetap relevan dan andal pada konteks layanan publik skala lokal yang memiliki keterbatasan data dan sumber daya komputasi. Dominasi sentimen positif sebesar 63% menunjukkan bahwa layanan Kantor Desa Linggang Bigung telah memenuhi ekspektasi mayoritas masyarakat, terutama pada aspek kecepatan pelayanan, keramahan petugas, dan kejelasan prosedur administrasi, sementara sentimen negatif yang bersifat moderat terutama berkaitan dengan keterbatasan fasilitas fisik dan pengelolaan waktu tunggu. Temuan ini menegaskan bahwa kualitas interaksi aparatur desa merupakan faktor utama pembentuk kepuasan layanan, sekaligus menunjukkan bahwa perbaikan operasional yang terfokus dapat meningkatkan kualitas layanan secara menyeluruh. Secara praktis, hasil penelitian ini dapat dimanfaatkan oleh pemerintah desa sebagai instrumen evaluasi layanan yang transparan, efisien, dan berkelanjutan, dengan prioritas perbaikan pada infrastruktur fisik dan efisiensi waktu layanan tanpa mengabaikan penguatan kualitas sumber daya manusia. Ke depan, pengembangan leksikon khusus konteks administrasi desa, perluasan cakupan data lintas desa, penerapan studi longitudinal, serta integrasi pendekatan lexicon-based dengan metode pembelajaran mesin direkomendasikan untuk meningkatkan sensitivitas analisis dan memperluas generalisasi temuan.

6. UCAPAN TERIMA KASIH

Penulis menyampaikan terima kasih dan apresiasi kepada Kantor Desa Linggang Bigung dan semua partisipan yang mengambil bagian dalam studi ini dan memberikan partisipan tersebut rasa hormat dan memiliki pemahaman yang baik tentang kualitas layanan desa. Kami mengakui STMIK Widya Cipta Dharma untuk dukungan yang telah diberikan kepada kami untuk melaksanakan penelitian. Kami mengirimkan terima kasih terhangat kami kepada mereka yang secara kritis membaca artikel kami



dan memberi kami umpan balik dalam perencanaan penelitian dan tahap penulisan naskah.

DAFTAR PUSTAKA:

- [1] M. Kumar, L. Khan, and H. T. Chang, "Evolving techniques in sentiment analysis: a comprehensive review," 2025, *PeerJ Inc.* doi: 10.7717/PEERJ-CS.2592.
- [2] A. A. N. Risal, Fathahillah, and D. R. A. Sulaiman, "Classification of Sentiment Analysis and Community Opinion Modeling Topics for Application of ICT in Government Operations," *Int. J. Environ. Eng. Educ.*, vol. 5, no. 1, pp. 35–44, Apr. 2023, doi: 10.55151/ijeedu.v5i1.99.
- [3] J. R. Jim, M. A. R. Talukder, P. Malakar, M. M. Kabir, K. Nur, and M. F. Mridha, "Recent advancements and challenges of NLP-based sentiment analysis: A state-of-the-art review," Mar. 01, 2024, *Elsevier Ltd.* doi: 10.1016/j.nlp.2024.100059.
- [4] A. Ligthart, C. Catal, and B. Tekinerdogan, "Systematic reviews in sentiment analysis: a tertiary study," *Artif. Intell. Rev.*, vol. 54, no. 7, pp. 4997–5053, Oct. 2021, doi: 10.1007/s10462-021-09973-3.
- [5] Z. Yao, J. Yang, J. Liu, M. Keith, and C. H. Guan, "Comparing tweet sentiments in megacities using machine learning techniques: In the midst of COVID-19," *Cities*, vol. 116, Sep. 2021, doi: 10.1016/j.cities.2021.103273.
- [6] F. Alshuaier, A. Areshey, and J. Poon, "Applications and Enhancement of Document-Based Sentiment Analysis in Deep learning Methods: Systematic Literature Review," Sep. 01, 2022, *Elsevier B.V.* doi: 10.1016/j.iswa.2022.200090.
- [7] G. Kaur and A. Sharma, "A deep learning-based model using hybrid feature extraction approach for consumer sentiment analysis," *J. Big Data*, vol. 10, no. 1, Dec. 2023, doi: 10.1186/s40537-022-00680-6.
- [8] N. A. M. Razali *et al.*, "Opinion mining for national security: techniques, domain applications, challenges and research opportunities," *J. Big Data*, vol. 8, no. 1, Dec. 2021, doi: 10.1186/s40537-021-00536-5.
- [9] A. S. Talaat, "Sentiment analysis classification system using hybrid BERT models," *J. Big Data*, vol. 10, no. 1, Dec. 2023, doi: 10.1186/s40537-023-00781-w.
- [10] M. Wankhade, A. C. S. Rao, and C. Kulkarni, "A survey on sentiment analysis methods, applications, and challenges," *Artif. Intell. Rev.*, vol. 55, no. 7, pp. 5731–5780, Oct. 2022, doi: 10.1007/s10462-022-10144-1.
- [11] A. M. Van Der Veen and E. Bleich, "The advantages of lexicon-based sentiment analysis in an age of machine learning," *PLoS One*, vol. 20, no. 1 January, Jan. 2025, doi: 10.1371/journal.pone.0313092.
- [12] P. Wang, J. Li, and J. Hou, "S2SAN: A sentence-to-sentence attention network for sentiment analysis of online reviews," *Decis. Support Syst.*, vol. 149, Oct. 2021, doi: 10.1016/j.dss.2021.113603.
- [13] L. Xu and W. Wang, "Improving aspect-based sentiment analysis with contrastive learning," *Nat. Lang. Process. J.*, vol. 3, p. 100009, Jun. 2023, doi: 10.1016/j.nlp.2023.100009.
- [14] P. Mukherjee, Y. Badr, S. Doppalapudi, S. M. Srinivasan, R. S. Sangwan, and R. Sharma, "Effect of Negation in Sentences on Sentiment Analysis and Polarity Detection," in *Procedia Computer Science*, Elsevier B.V., 2021, pp. 370–379. doi: 10.1016/j.procs.2021.05.038.
- [15] F. B. Matheus Araújo, Adriano Pereira, "A COMPARATIVE STUDY OF MACHINE TRANSLATION FOR MULTILINGUAL SENTENCE-LEVEL SENTIMENT ANALYSIS." doi: <https://doi.org/10.1016/j.ins.2019.10.031>.
- [16] M. A. Saddam, E. K. Dewantara, and A. Solichin, "Sentiment Analysis of Flood Disaster Management in Jakarta on Twitter Using Support Vector Machines," *Sinkron*, vol. 8, no. 1, pp. 470–479, Jan. 2023, doi: 10.33395/sinkron.v8i1.12063.
- [17] C. Haynes *et al.*, "Automatic Classification of National Health Service Feedback," *Mathematics*, vol. 10, no. 6, Mar. 2022, doi: 10.3390/math10060983.
- [18] H. T. Duong and T. A. Nguyen-Thi, "A review: preprocessing techniques and data augmentation for sentiment analysis," *Comput. Soc. Networks*, vol. 8, no. 1, Dec. 2021, doi: 10.1186/s40649-020-00080-x.
- [19] M. Sunitha, T. Adilakshmi, and V. S. Vaishnavi, "Movie Recommendation and Sentiment Analysis using Deep Learning



- [20] Algorithms,” 2024. [Online]. Available: <https://www.jisem-journal.com/>
- [21] R. Kowalski, M. Esteve, and S. Jankin Mikhaylov, “Improving public services by mining citizen feedback: An application of natural language processing,” *Public Adm.*, vol. 98, no. 4, pp. 1011–1026, Dec. 2020, doi: 10.1111/padm.12656.
- [22] R. Srivastava, P. K. Bharti, and P. Verma, “Comparative Analysis of Lexicon and Machine Learning Approach for Sentiment Analysis.” [Online]. Available: www.ijacsa.thesai.org
- [23] M. S. Assyifa *et al.*, “Sentiment Analysis of STMIK Widya Cipta Dharma Using a Lexicon-Based Approach Analisis Sentimen Terhadap STMIK Widya Cipta Dharma Menggunakan Pendekatan Lexicon,” pp. 1–7, 2021.
- C. H. Hill, J. E. Fresneda, and M. Anandarajan, “The wisdom of the lexicon crowds: leveraging on decades of lexicon-based sentiment analysis for improved results,” *J. Big Data*, vol. 12, no. 1, Dec. 2025, doi: 10.1186/s40537-025-01186-7.