

ANALISIS SENTIMEN ULASAN MEDIA SOSIAL UMKM KULINER DENGAN PENDEKATAN LEXICON-BASED DAN KOSAKATA KHUSUS

Setyo Wahyu Saputro¹, Friska Abadi², Radityo Adi Nugroho³

^{1,2,3} Program Studi Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Lambung Mangkurat

¹setyo.saputro@ulm.ac.id, ²friska.abadi@ulm.ac.id, ³radityo.adi@ulm.ac.id

Abstrak

UMKM kuliner di Kalimantan Selatan memanfaatkan media sosial sebagai sarana utama untuk mengetahui opini pelanggan, namun jumlah komentar yang sangat besar menyulitkan pelaku usaha untuk menelaahnya secara manual. Kondisi ini menegaskan perlunya pendekatan analisis sentimen yang mampu mengolah data ulasan secara efisien serta sesuai dengan karakteristik bahasa lokal. Penelitian ini bertujuan mengembangkan metode analisis sentimen berbasis *lexicon* yang diperkaya dengan kosakata domain-spesifik kuliner dan bahasa Banjar agar hasil klasifikasi lebih akurat dan kontekstual. Data penelitian diperoleh dari 3.500 komentar publik di Instagram dan TikTok. Tahap preprocessing mencakup case folding, pembersihan karakter khusus, tokenisasi, *stopword removal*, normalisasi, dan *stemming*. Selanjutnya, *InSet Lexicon* disempurnakan melalui penyuntikan kosakata baru serta penyesuaian bobot kata sesuai konteks kuliner lokal. Hasil analisis menunjukkan distribusi sentimen terdiri dari 2.050 komentar positif (58,57%), 934 komentar netral (26,69%), dan 516 komentar negatif (14,74%). Evaluasi menunjukkan peningkatan akurasi signifikan setelah perluasan *lexicon*, yaitu 93,49% untuk sentimen negatif, 94,64% untuk netral, dan 96,94% untuk positif, dibandingkan akurasi awal yang berkisar antara 51–73%. Temuan ini membuktikan bahwa pengayaan *lexicon* menggunakan kosakata lokal dan domain-spesifik secara substansial meningkatkan performa analisis sentimen. Pendekatan ini memberikan solusi praktis dan terjangkau bagi UMKM untuk memahami opini pelanggan secara lebih representatif, serta dapat dimanfaatkan dalam pengambilan keputusan strategis dan perbaikan kualitas layanan maupun promosi produk kuliner.

Kata kunci: Analisis sentimen, *lexicon-based*, kosakata domain-spesifik, media sosial, UMKM kuliner

1. Pendahuluan

Usaha Mikro, Kecil, dan Menengah (UMKM) merupakan salah satu potensi utama di Provinsi Kalimantan Selatan. Berdasarkan data dari Badan Pusat Statistik (BPS) Kalimantan Selatan, data yang dipublikasi tahun 2023 berdasarkan data tahun 2013 hingga 2021 terdapat 64.528 usaha mikro dan 2.842 usaha kecil (Badan Pusat Statistik, 2024). Pada tahun 2023, Pemerintah Provinsi Kalimantan Selatan melakukan pendataan ulang terhadap sekitar 375.000 UMKM yang tersebar di 13 kabupaten/kota, dengan bantuan dari Kementerian Koperasi dan Usaha Kecil dan Menengah Republik Indonesia. Data menunjukkan bahwa di lima kabupaten/kota, yaitu Kota Banjarmasin, Kabupaten Banjar, Kabupaten Hulu Sungai Selatan, dan Kabupaten Tanah Laut. Dinas Koperasi dan UMKM Provinsi Kalimantan Selatan memperkirakan bahwa jumlah total UMKM di seluruh provinsi dapat mencapai 500.000 unit (Dinas Koperasi, 2024).

Sektor kuliner sebagai salah satu sektor UMKM tertinggi di Kalimantan Selatan, perlu mendapatkan perhatian khusus dikarenakan memiliki karakteristik unik dari produknya (Mahdi Hidayatullah & Hastin Umi Anisah, 2024). Produk kuliner memiliki sifat produk yang tergolong cepat rusak dan lebih sensitif terhadap selera dan tren pasar.

Dalam menghadapi persaingan yang semakin ketat, UMKM kuliner perlu memahami opini dan preferensi pelanggan untuk meningkatkan kualitas produk dan layanan mereka. Salah satu sumber utama informasi terkait opini pelanggan adalah media sosial, yang memungkinkan pelanggan untuk memberikan ulasan, komentar, dan rekomendasi secara langsung (Fadel Sasongko et al., n.d.).

Media sosial telah menjadi platform utama bagi konsumen untuk mengekspresikan pendapat mereka terhadap suatu produk atau layanan. Ulasan pelanggan di platform seperti Instagram, Twitter, TikTok, dan Google Reviews dapat memberikan wawasan berharga tentang kepuasan pelanggan, tren pasar, serta kelemahan dan keunggulan produk UMKM kuliner (Adhityas et al., n.d.). Namun, besarnya volume data dari media sosial menyebabkan kesulitan dalam mengolah informasi tersebut secara manual, diperlukan suatu pendekatan analitis untuk mengekstraksi opini pelanggan dengan cara yang lebih sistematis dan efisien (Nadia L Kamila, 2023).

Pendekatan berbasis *Lexicon* (*lexicon-based*) dalam analisis sentimen adalah metode yang menggunakan kamus kata (*Lexicon*) untuk menentukan polaritas sentimen dalam teks (Fauziah et al., 2021). Kamus ini berisi daftar kata dengan skor polaritas yang menunjukkan apakah kata tersebut memiliki sentimen positif, negatif, atau netral. Setiap

kata dalam teks dianalisis dengan mencocokkannya dengan entri dalam kamus untuk menentukan sentimen keseluruhan teks tersebut (Bagas et al., 2024). Lexicon-Based merupakan salah satu metode ekstraksi fitur yang dapat meningkatkan performansi fitur. Metode ini bekerja dengan cara membuat kamus berupa kata terlebih dahulu, lalu kata-kata tersebut digunakan untuk mengidentifikasi apakah kata-kata tersebut mengandung opini atau tidak (Prayuda & Permata, 2022).

Penelitian ini pendekatan berbasis kamus Lexicon menggunakan *InSet Lexicon* yang dibuat oleh Fajri Koto dan Gemala Y Rahmaningtyas (Koto & Rahmaningtyas, 2017), dan juga memanfaatkan kosakata khusus kuliner yang dimaksudkan untuk memberikan informasi spesifik mengenai kata-kata yang sering digunakan oleh konsumen dalam memberikan respons terhadap produk kuliner berikut juga dengan bahasa lokal Kalimantan Selatan, bahasa Banjar. Dengan menerapkan metode ini, UMKM kuliner dapat mengenali aspek-aspek yang mendapatkan respons positif maupun negatif dari pelanggan, yang selanjutnya dapat dijadikan landasan dalam peningkatan kualitas produk dan strategi pemasaran (Bhowmick & Seetharaman, 2023).

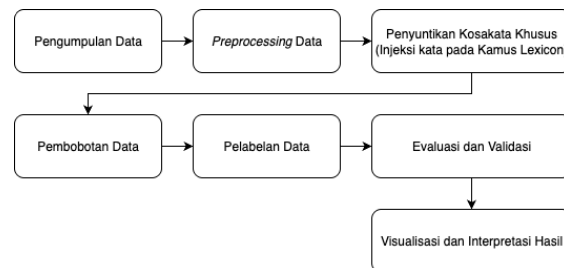
Urgensi penelitian ini berlandaskan pada kenyataan bahwa UMKM kuliner di Kalimantan Selatan memiliki potensi besar, namun masih menghadapi tantangan dalam memahami opini pelanggan yang tersebar luas di media sosial. Pelaku UMKM tidak memiliki kemampuan atau alat yang memadai untuk menganalisis ribuan ulasan secara manual, sedangkan ulasan tersebut diperlukan untuk mengetahui kekuatan dan kelemahan produknya (McCloskey et al., 2024). Banyak perusahaan juga menemukan wawasan dari ulasan pelanggan untuk merumuskan rencana strategis berdasarkan pola preferensi pelanggan yang memiliki dampak signifikan terhadap keuntungan keseluruhan perusahaan (Subhashini et al., 2021).

Penelitian ini melakukan analisis sentimen berbasis *lexicon* dengan kosakata khas kuliner lokal, guna membantu UMKM memahami opini konsumen melalui pendekatan berbasis data. Dirahapkan tidak hanya menjawab kebutuhan praktis pelaku UMKM dalam pengambilan keputusan, tetapi juga mengisi kesenjangan ilmiah dalam pengembangan analisis sentimen berbasis *lexicon* yang relevan secara kontekstual. Pendekatan ini menawarkan solusi yang praktis dan terjangkau, tanpa harus bergantung pada teknologi machine learning yang kompleks dan memerlukan sumber daya besar.

2. Metode

Penelitian dilakukan dengan sejumlah tahapan sistematis dan berurutan, sebagaimana ditampilkan pada Gambar 1. Tahapan penelitian diantaranya pengumpulan data, *preprocessing* data, penyuntikan

kosakata, pembobotan data, pelabelan data, evaluasi & validasi, visualisasi & interpretasi hasil



Gambar 1. Diagram proses metodologi

Sebagai metode pembandingan (*baseline*), penelitian ini menggunakan pendekatan analisis sentimen berbasis *lexicon* dengan *Inset Lexicon* standar tanpa penambahan kosakata domain-spesifik. *Baseline* ini merepresentasikan pendekatan *lexicon-based* umum yang digunakan pada analisis sentimen bahasa Indonesia. Perbandingan kinerja antara *baseline* dan metode usulan dilakukan pada tahap evaluasi untuk menilai dampak penyuntikan kosakata kuliner dan bahasa Banjar terhadap performa klasifikasi.

2.1 Pengumpulan Data

Tahap awal dalam penelitian ini adalah pengumpulan data dengan melakukan scraping data komentar publik di Instagram dan TikTok. Pengambilan data dilakukan melalui API dengan memperhatikan aspek etika, yaitu hanya memanfaatkan komentar yang bersifat terbuka dan tidak melibatkan informasi pribadi pengguna untuk menjaga privasi, keamanan, ekspektasi pengguna, dan meminimalkan dampak negatif baik terhadap individu maupun platform (Brown et al., 2024). Dari tahap ini diperoleh dataset mentah sebanyak 3.500 komentar, beberapa komentar tersaji pada Tabel 1.

Tabel 1. Dataset Komentar

Komentar
nasi kuning non kolestrol kaka
@w3we3 non php kakaaa
Bagus kak 😊
@anisakeke mandai mba..dibuat dri kulit cempedak 😊😊😊
Ada di gofood gak Kaka?
Ga salah donk, wong peminatnya banyak 😊 (bantujawab)
Enak enak semua makanannya di sana
Dasar nyaman pang makanannya
Yg ini ready gk hari ini?
.....

2.2 Preprocessing Data

Tahap selanjutnya adalah *data preprocessing*. Tahap *preprocessing* data dilakukan untuk menyiapkan teks agar siap dianalisis dengan cara membersihkan dan menstandarkan isi data. Proses ini mencakup penyamaan huruf menjadi bentuk kecil

(*case folding*), penghapusan elemen yang tidak relevan seperti simbol atau tanda baca, pemecahan teks menjadi kata-kata (*tokenisasi*), penghilangan kata umum yang tidak memiliki makna penting (*stopword removal*), penyeragaman bentuk kata agar sesuai dengan kaidah baku (*normalisasi*), serta pengembalian kata berimbuhan ke bentuk dasar (*stemming*). Seluruh langkah ini bertujuan untuk menghasilkan data teks yang bersih, konsisten, dan siap digunakan pada tahap analisis berikutnya (Budiarto Hadiprakoso et al., n.d.).

a. Case Folding

Proses ini menyeragamkan semua huruf komentar menjadi huruf kecil (lowercase). Bertujuan agar kata yang sama dengan perbedaan kapitalisasi, seperti “Enak” dan “enak” tetap diperlakukan sama.

Tabel 2. Hasil *Case Folding* Komentar

Komentar	Case Folding
Ada di gofood gak Kaka?	ada di gofood gak kaka?
Ga salah donk, wong peminatnya banyak 😊 (bantujawab)	ga salah donk, wong peminatnya banyak 😊 (bantujawab)
Enak enak semua makanannya di sana	enak enak semua makanannya di sana
Dasar nyaman pang makanannya	dasar nyaman pang makanannya
Yg ini ready gk hari ini?	yg ini ready gk hari ini?
.....	

b. Pembersihan Karakter Khusus

Proses selanjutnya adalah menghilangkan elemen yang tidak relevan terhadap teks, seperti URL, mention, hastag, emoji, dan double spasi. Tahapan ini memastikan hanya teks utama yang dianalisis sehingga hasil klasifikasi sentimen lebih akurat. Hasil pembersihan dari karakter khusus dapat dilihat pada Tabel 3.

Tabel 3. Hasil *Case Folding* Komentar

Komentar	Cleaning
Ada di gofood gak Kaka?	ada di gofood gak kaka
Ga salah donk, wong peminatnya banyak 😊 (bantujawab)	ga salah donk wong peminatnya banyak bantujawab
Enak enak semua makanannya di sana	enak enak semua makanannya di sana
Dasar nyaman pang makanannya	dasar nyaman pang makanannya
Yg ini ready gk hari ini?	yg ini ready gk hari ini
.....	

c. Tokenisasi

Proses selanjutnya adalah memecah komentar menjadi potongan kata (*tokens*) agar dapat dianalisis secara terpisah. Proses digunakan untuk mengidentifikasi kata kunci yang memiliki bobot sentimen dalam komentar. Hasil tokenisasi komentar dapat dilihat pada Tabel 4.

Tabel 4. Hasil Tokenisasi Komentar

Komentar	Tokenisasi
Ada di gofood gak Kaka?	['ada', 'di', 'gofood', 'gak', 'kaka']
Ga salah donk, wong peminatnya banyak 😊 (bantujawab)	['ga', 'salah', 'donk', 'wong', 'peminatnya', 'banyak', 'bantujawab']
Enak enak semua makanannya di sana	['enak', 'enak', 'semua', 'makanannya', 'di', 'sana']
Dasar nyaman pang makanannya	['dasar', 'nyaman', 'pang', 'makanannya']
Yg ini ready gk hari ini?	['yg', 'ini', 'ready', 'gk', 'hari', 'ini']
.....	

d. Filtering/ Stopword Removal

Tahapan filtering yang digunakan untuk mengambil kata-kata yang penting. Kata umum yang biasanya muncul dan tidak memiliki makna disebut dengan *stopword*. Penghilangan *stopword* ini dapat mengurangi waktu pemrosesan. Hasil filtering kata dapat dilihat pada Tabel 5.

Tabel 5. Hasil Tokenisasi Komentar

Komentar	Tokenisasi
Ada di gofood gak Kaka?	['ada', 'gofood', 'gak', 'kaka']
Ga salah donk, wong peminatnya banyak 😊 (bantujawab)	['ga', 'salah', 'peminatnya', 'banyak', 'bantujawab']
Enak enak semua makanannya di sana	['enak', 'enak', 'semua', 'makanannya']
Dasar nyaman pang makanannya	['dasar', 'nyaman', 'makanannya']
Yg ini ready gk hari ini?	['ready', 'gk']
.....	

e. Normalisasi Teks

Komentar berisi singkatan, slang, atau penulisan tidak baku, maka dilakukan proses normalisasi. Misalnya kata “gk”, “ga”, “gak” dinormalisasi menjadi “tidak”. Hasil normalisasi teks dapat dilihat pada Tabel 6.

Tabel 6. Hasil Normalisasi Teks

Komentar	Tokenisasi
Ada di gofood gak Kaka?	['ada', 'gofood', 'tidak', 'kaka']
Ga salah donk, wong peminatnya banyak 😊 (bantujawab)	['tidak', 'salah', 'peminatnya', 'banyak', 'bantujawab']
Enak enak semua makanannya di sana	['enak', 'enak', 'semua', 'makanannya']
Dasar nyaman pang makanannya	['dasar', 'nyaman', 'makanannya']
Yg ini ready gk hari ini?	['ready', 'tidak']
.....	

f. Stemming

Kata dalam komentar yang memiliki imbuhan dikembalikan ke bentuk dasar, misalnya kata “makanannya” menjadi “makan”. Hasil *stemming* komentar dapat dilihat pada Tabel 7.

Tabel 7. Hasil *Stemming* Teks

Komentar	Tokenisasi
Ada di gofood gak Kaka?	['ada', 'gofood', 'tidak', 'kaka']
Ga salah donk, wong peminatnya banyak 😊 (bantu jawab)	['tidak', 'salah', 'minat', 'banyak', 'bantu jawab']
Enak enak semua makanannya di sana	['enak', 'enak', 'semua', 'makan']
Dasar nyaman pang makanannya	['dasar', 'nyaman', 'makan']
Yg ini ready gk hari ini?	['ready', 'tidak']
.....	

2.3 Penyuntikan Kosakata Khusus

Metode klasifikasi *lexicon* memerlukan kamus kata yang berisi daftar kata positif dan negatif beserta bobotnya. Karena terdapat banyak kata baru terutama bahasa suku banjar, singkatan, maupun istilah khas pengguna media sosial yang belum tercakup dalam kamus, peneliti melakukan penyuntikan kata baru ke dalam *lexicon*. Kamus *lexicon* yang diperluas dengan kata-domain spesifik dari korpus domain dapat meningkatkan performa analisis sentimen pada data domain tersebut (Shaukat et al., 2020). Hasil skor komentar setelah penambahan kosakata dapat dilihat pada Tabel 8.

Tabel 8. Tabel Skor Komentar dengan Penambahan Kosakata pada Inset Lexicon

Komentar	Skor Sebelum	Skor Sesudah
['ada', 'gofood', 'tidak', 'kaka']	-3	4
['tidak', 'salah', 'minat', 'banyak', 'bantu jawab']	4	4
['enak', 'enak', 'semua', 'makan']	12	12
['dasar', 'nyaman', 'makan']	-1	5
['ready', 'tidak']	-2	1
.....		

2.4 Ekstraksi Fitur

Ekstraksi fitur dalam analisis sentimen berbasis *lexicon* dilakukan untuk secara otomatis mengubah teks menjadi representasi numerik sehingga sentimen dapat dianalisis dengan efisien, mengurangi ketergantungan pada pengkodean manual yang mahal dan memakan waktu, serta memungkinkan pemrosesan data teks dalam jumlah besar (Cero et al., 2024).

Setiap komentar direpresentasikan ke dalam bentuk fitur numerik berdasarkan keberadaan kata-kata dalam *lexicon*. Bobot sentimen dihitung dengan menjumlahkan skor kata positif dan negatif yang terdapat dalam komentar. Jika komentar mengandung lebih banyak kata positif, maka total skor bernilai positif, begitu pula sebaliknya untuk komentar negatif. Sedangkan jika skor berimbang atau nol, komentar dikategorikan sebagai netral. Skor hasil ekstraksi fitur dapat dilihat pada Tabel 9.

Tabel 9. Skor Komentar dengan Penambahan Kosakata pada Inset Lexicon

Komentar	Skor
['ada', 'gofood', 'tidak', 'kaka']	4
['tidak', 'salah', 'minat', 'banyak', 'bantu jawab']	4
['enak', 'enak', 'semua', 'makan']	12
['dasar', 'nyaman', 'makan']	5
['ready', 'tidak']	1
.....	

2.5 Pelabelan data dan Validasi

Proses pelabelan dilakukan secara semi otomatis dengan pendekatan *lexicon-based*. Setiap komentar diberi label berdasarkan skor polaritas, dihasilkan dari *Inset Lexicon* yang diperluas dengan kosakata domain kuliner dan bahasa Banjar. Untuk memastikan kebenaran hasil label, dilakukan validasi manual terhadap 10% komentar yang dipilih secara acak dari dataset dengan tiga kelas sentimen yaitu positif, negatif, dan netral. Validasi dilakukan oleh tiga penilai independen yang memiliki latar belakang ilmu komputer serta memahami konteks kuliner dan bahasa lokal Banjar. Penentuan label akhir menggunakan *majority voting*, jika terdapat ketidaksepakatan label antar penilai, label ditentukan berdasarkan suara terbanyak.

Mayoritas komentar memperoleh kesepakatan label pada penilaian awal, sementara perbedaan penilaian umumnya terjadi pada komentar yang bersifat ambigu. Hasil validasi menunjukkan bahwa sebagian besar komentar memperoleh kesepakatan label yang konsisten antar penilai, sehingga data hasil pelabelan dapat digunakan sebagai acuan evaluasi sistem.

2.6 Klasifikasi

Klasifikasi sentimen dilakukan berdasarkan skor polaritas yang diperoleh dari tahap ekstraksi fitur.

Jika skor sentimen > 0 maka komentar diklasifikasikan sebagai positif.

Jika skor sentimen < 0 maka komentar diklasifikasikan sebagai negatif.

Jika skor sentimen $= 0$ maka komentar diklasifikasikan sebagai netral.

Hasil sementara total komentar positif 2033, netral 1117, dan negatif 350.

Berdasarkan proses klasifikasi tersebut, diperoleh distribusi sentimen yang selanjutnya digunakan dalam tahap evaluasi kinerja sistem.

Tabel 10. Klasifikasi Positif Negatif Komentar

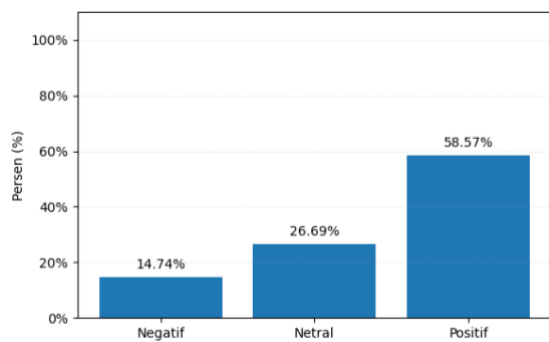
Komentar	Skor Sebelum	Skor Sesudah
['ada', 'gofood', 'tidak', 'kaka']	4	Positif
['tidak', 'salah', 'minat', 'banyak', 'bantu jawab']	4	Positif
['enak', 'enak', 'semua', 'makan']	12	Positif

['dasar', 'nyaman', 'makan']	5	Positif
['ready', 'tidak']	1	Positif
.....		

3. Hasil dan Pembahasan

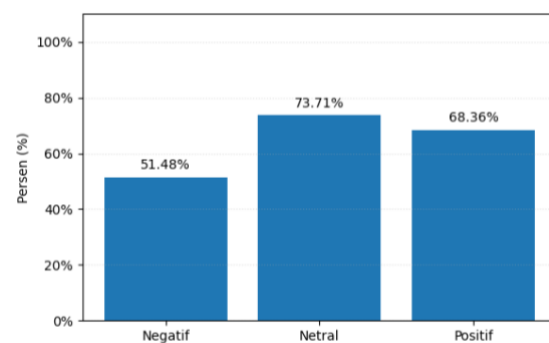
Penelitian ini mengusulkan pendekatan analisis sentimen berbasis *lexicon* yang disesuaikan dengan konteks kuliner lokal. Data diambil dari 3.500 komentar publik di Instagram dan TikTok. Setelah melalui tahapan preprocessing (case folding, cleaning, tokenisasi, stopwords removal, normalisasi, dan stemming), data dianalisis menggunakan InSet *Lexicon* yang telah diperluas dengan kosakata khas kuliner dan bahasa Banjar.

Distribusi sentimen pada Gambar 2 menunjukkan komentar positif sebanyak 2.050 data (58,57%), komentar netral sebanyak 934 (26,69%) dan komentar negatif sebanyak 516 (14,74%).



Gambar 2. Distribusi Hasil Sentimen

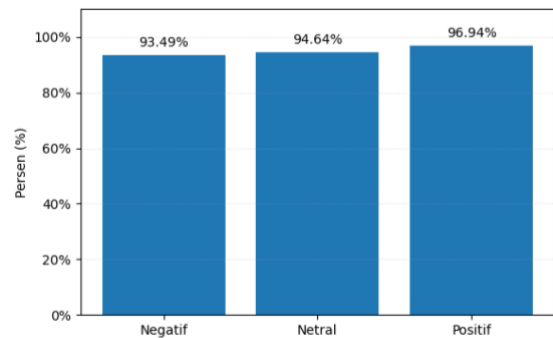
Hasil ini menunjukkan bahwa metode *lexicon* dengan penyuntikan kosakata kuliner dan kosakata lokal meningkatkan performa analisis signifikan dibandingkan penggunaan InSet *Lexicon* standar. Pada Gambar 3 ditunjukkan evaluasi awal tanpa penyuntikan kosakata akurasi mencapai sekitar 68,36% (positif) 73,71% (netral), dan 51,48% (negatif). Setelah penyesuaian bobot dan penambahan kata baru seperti “nyaman pang”, “mantap”, dan “gasan” performa meningkat drastis pada seluruh kategori.



Gambar 3. Akurasi Sebelum Perbaikan dan Suntikan Kosakata Baru

Setelah dilakukan berbagai perbaikan pada kamus dan melakukan penambahan kosakata baru

yang relevan, performa model meningkat secara signifikan. Proses evaluasi dilakukan menggunakan confusion matrix, akurasi pada kategori negatif naik hingga 93,49%, kategori netral mencapai 94,64%, dan kategori positif mencapai 96,94% (Gambar 4). Peningkatan ini memperlihatkan efektivitas perbaikan yang dilakukan, sekaligus menegaskan pentingnya penyesuaian kamus agar selaras dengan konteks data yang dianalisis.



Gambar 4. Akurasi Sesudah Perbaikan dan Suntikan Kosakata Baru

Evaluasi kinerja dilakukan dengan membandingkan hasil klasifikasi metode usulan terhadap metode *baseline* menggunakan confusion matrix dan akurasi per kelas sentimen. Pada *baseline*, performa klasifikasi masih terbatas terutama pada kelas negatif dan netral, yang disebabkan oleh keterbatasan cakupan kosakata dalam Inset *lexicon* standar. Setelah dilakukan penyuntikan domain spesifik, akurasi klasifikasi meningkat secara konsisten, menunjukkan bahwa penyesuaian *lexicon* terhadap konteks domain berkontribusi langsung terhadap peningkatan performa sistem.

Peningkatan ini memperkuat temuan (Bagas et al., 2024) yang menyatakan bahwa penyesuaian *lexicon* sesuai konteks domain dapat meningkatkan akurasi analisis sentimen secara signifikan. Hasil penelitian juga selaras dengan studi (Fadel Sasongko et al., n.d.) yang menemukan bahwa konteks kuliner lokal memiliki karakteristik linguistik unik yang tidak sepenuhnya tercakup dalam kamus sentimen umum.

4. Kesimpulan

Penelitian ini berhasil menerapkan metode *lexicon*-based dengan penyuntikan kosakata domain-spesifik kuliner dan bahasa Banjar untuk menganalisis sentimen ulasan konsumen terhadap produk UMKM kuliner di media sosial. Hasil menunjukkan bahwa sentimen positif mendominasi sebesar 58,57%, dengan akurasi model mencapai 95,83% setelah penyesuaian kamus dan bobot kata.

Temuan ini membuktikan bahwa penyesuaian *lexicon* sesuai konteks lokal secara signifikan meningkatkan ketepatan analisis dan mampu menggambarkan opini publik secara lebih representatif. Pendekatan ini dapat diterapkan oleh pelaku UMKM sebagai dasar pengambilan keputusan

berbasis data untuk meningkatkan kualitas layanan dan strategi promosi.

Penelitian ini memiliki beberapa keterbatasan. Pendekatan *lexicon-based* belum sepenuhnya mampu menangkap makna kontekstual yang muncul pada komentar di media sosial. Selain itu, penyusunan kosakata domain-spesifik masih memerlukan intervensi manual sehingga perlu diperbaharui secara mengikuti perkembangan bahasa.

Untuk penelitian selanjutnya, disarankan untuk memperluas pengembangan dengan melakukan penyempurnaan kosakata domain-spesifik agar lebih adaptif terhadap perubahan tren bahasa di media sosial dan konteks lokal. Selain itu, pengembangan dashboard web interaktif diperlukan guna memvisualisasikan distribusi sentimen secara real-time sehingga hasil analisis dapat dimanfaatkan langsung oleh pelaku UMKM dalam pengambilan keputusan. Pengayaan dataset dari berbagai platform media sosial juga diharapkan dapat meningkatkan keberagaman data dan memberikan hasil analisis yang lebih komprehensif terhadap opini konsumen.

Daftar Pustaka:

- Adhityas, A. S., Satya Nugraha, S., & Syuhada, M. A. (n.d.). Analisis Sentimen Ulasan Pelanggan pada Kafe Gamat Bendungan Hilir Menggunakan Algoritma Naive bayes. In *Purwokerto Kidul, Kec. Purwokerto Sel* (Vol. 4, Issue 1).
- Badan Pusat Statistik. (2024). *Jumlah Perusahaan Industri Skala Mikro dan Kecil Menurut Provinsi (Unit)*, 2024.
- Bagas, M., Putra, D., & Setiawan, E. (2024). Metode Lexicon Based untuk Analisis Sentimen Pengguna Twitter terhadap Kinerja ISP (Studi Kasus: IndiHome, Biznet, MyRepublic). In *Jurnal Mahasiswa Teknik Informatika* (Vol. 8, Issue 6). <http://dev.twitter.com>
- Bhowmick, A., & Seetharaman, A. (2023). Impact of product quality on customer satisfaction: A Systematic Literature Review. *ACM International Conference Proceeding Series*, 93–99. <https://doi.org/10.1145/3603421.3603434>
- Brown, M. A., Gruen, A., Maldoff, G., Messing, S., Sanderson, Z., & Zimmer, M. (2024). *Web Scraping for Research: Legal, Ethical, Institutional, and Scientific Considerations*. <http://arxiv.org/abs/2410.23432>
- Budiarto Hadiprakoso, R., Setiawan, H., Novita Yasa, R., & Siber dan Sandi Negara, P. (n.d.). *Text Preprocessing for Optimal Accuracy in Indonesian Sentiment Analysis Using a Deep Learning Model with Word Embedding*.
- Cero, I., Luo, J., & Falligant, J. M. (2024). Lexicon-Based Sentiment Analysis in Behavioral Research. *Perspectives on Behavior Science*, 47(1), 283–310. <https://doi.org/10.1007/s40614-023-00394-x>
- Dinas Koperasi, U. K. dan M. P. K. S. (2024, October). *Jumlah Usaha Kecil, Mikro dan Menengah Provinsi Kalimantan Selatan*.
- Fadel Sasongko, M., Hayati, N., Informasi, S., & Nasional, U. (n.d.). *Analisis Sentimen Opini Masyarakat Terhadap Kuliner DKI Jakarta Dengan Metode Naïve Baiyes dan Support Vector Machine*.
- Fauziah, Y., Yuwono, B., & Aribowo, A. S. (2021). Lexicon Based Sentiment Analysis in Indonesia Languages : A Systematic Literature Review. *RSF Conference Series: Engineering and Technology*, 1(1), 363–367. <https://doi.org/10.31098/cset.v1i1.397>
- Koto, F., & Rahmaningtyas, G. Y. (2017). Inset lexicon: Evaluation of a word list for Indonesian sentiment analysis in microblogs. *2017 International Conference on Asian Language Processing (IALP)*, 391–394. <https://doi.org/10.1109/IALP.2017.8300625>
- Mahdi Hidayatullah, & Hastin Umi Anisah. (2024). Adaptasi Pasar terhadap Tren Makanan Viral di Banjarmasin, Kalimantan Selatan: Peluang Bisnis bagi Wirausaha Kuliner Lokal. *Jurnal Kajian Dan Penalaran Ilmu Manajemen*, 2(4), 70–92. <https://doi.org/10.59031/jkpm.v2i4.505>
- McCloskey, B. J., LaCasse, P. M., & Cox, B. A. (2024). Natural language processing analysis of online reviews for small business: extracting insight from small corpora. *Annals of Operations Research*, 341(1), 295–312. <https://doi.org/10.1007/s10479-023-05816-2>
- Nadia L Kamila. (2023, September 7). *Memahami Opini Publik dengan Analisis Sentimen Media Sosial*. <https://Dibimbing.Id/Blog/Detail/Memahami-Opini-Publik-Dengan-Analisis-Sentimen-Media-Sosial>.
- Prayuda, M. W., & Permana, A. A. (2022). *Penerapan Metode Lexicon Based untuk Menganalisis Sentimen Terhadap Mudik Lebaran. 11*.
- Shaukat, K., A Hameed, I., Luo, S., Javed, I., Iqbal, F., Faisal, A., Masood, R., Usman, A., Shaukat, U., Hassan, R., Younas, A., Ali, S., & Adeem, G. (2020). Domain Specific Lexicon Generation through Sentiment Analysis. *International Journal of Emerging Technologies in Learning (IJET)*, 15(09), 190. <https://doi.org/10.3991/ijet.v15i09.13109>
- Subhashini, L. D. C. S., Li, Y., Zhang, J., Atukorale, A. S., & Wu, Y. (2021). Mining and classifying customer reviews: a survey. *Artificial Intelligence Review*, 54(8), 6343–6389. <https://doi.org/10.1007/s10462-021-09955-5>