Techno.Com, Vol. 25 No. 1, Februari 2026: 170 – 186

Analisis Sentimen Pelanggan Restoran Menggunakan Algoritma *Multinomial Logistic Regression*

Customer Sentiment Analysis for Restaurants Using Multinomial Logistic Regression Algorithm

Hajrianti[1], **Heliawati Hamrul**[2], **A. Amirul Asnan Cirua***[3]
Teknik Informatika, Universitas Sulawesi Barat
HajriantiRazann@gmail.com[1] *heliawatyhamrul@unsulbar.ac.id*[2],
*amirulasnancirua@unsulbar.ac.id**[3]
Received 9 January 2026; Revised 19 February 2026; Accepted 21 February 2026

**Corresponding author*

Abstrak - Industri kuliner di Indonesia terus mengalami perkembangan pesat, termasuk dengan hadirnya merek nasional seperti Mie Gacoan yang baru membuka cabang di Sulawesi Barat Mamuju. Kehadiran restoran tersebut memunculkan beragam opini dari masyarakat yang disampaikan melalui ulasan pelanggan. Penelitian ini bertujuan untuk melakukan analisis sentimen pelanggan terkait layanan restoran dengan penerapan algoritma *Multinomial Logistic Regression.* Data dikumpulkan melalui dua sumber, yaitu ulasan pelanggan dari *scraping* Google Maps dan kuesioner, dengan total sebanyak 773 data ulasan yang melalui tahap *preprocessing* teks dan pembobotan TF-IDF. Penelitian ini juga menerapkan pendekatan *Explainable AI* dengan metode *Local Interpretable Model-agnostic Explanations* (LIME) guna memperjelas kata-kata yang paling berpengaruh terhadap hasil klasifikasi model. Berdasarkan hasil pengujian, algoritma *Multinomial Logistic Regression* mampu memberikan performa klasifikasi dengan akurasi mencapai 86,5%, presisi 86,3%, recall 86,5%, dan F1-score 85,8% dalam mengidentifikasi opini pelanggan restoran. Temuan ini menunjukkan bahwa metode berbasis *Machine Learning* dan *Explainable AI* dapat memberikan analisis sentimen terhadap persepsi pelanggan, sehingga menjadi referensi penting bagi pengambilan keputusan strategis dalam peningkatan kualitas layanan.

Kata kunci - Sentimen Pelanggan, *Multinomial Logistic Regression*, TF-IDF, *Explainable AI,* LIME

Abstract *- The culinary industry in Indonesia continues to experience rapid development, including the presence of national brands such as Mie Gacoan, which recently opened a branch in Mamuju, West Sulawesi. The presence of this restaurant has generated various opinions from the public, expressed through customer reviews. This study aims to analyze customer sentiment regarding restaurant services using the Multinomial Logistic Regression algorithm. Data were collected from two sources: customer reviews from Google Maps scraping and questionnaires, with a total of 773 review data points that went through text preprocessing and TF-IDF weighting. This study also applied the Explainable AI approach with the Local Interpretable Model-agnostic Explanations (LIME) method to clarify the words that most influence the model's classification results. Based on the test results, the Multinomial Logistic Regression algorithm was able to provide classification performance with an accuracy of 86.5%, a precision of 86.3%, a recall of 86.5%, and an F1-score of 85.8% in identifying restaurant customer opinions. These findings indicate that Machine Learning and Explainable AI-based methods can provide sentiment analysis of customer perceptions, thus becoming an important reference for strategic decision-making in improving service quality.*

Keywords*: Customer Sentiment, Multinomial Logistic Regression, TF-IDF, Explainable AI, LIME*

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI : 10.62411/tc.v25i1.15532

1. PENDAHULUAN

Industri kuliner di Indonesia terus berkembang pesat seiring dengan meningkatnya minat masyarakat terhadap aktivitas makan di luar rumah. Sektor ini tidak hanya menjadi bagian dari gaya hidup modern, tetapi juga menjadi salah satu penopang ekonomi kreatif yang berkontribusi besar terhadap pertumbuhan daerah. Persaingan antar pelaku usaha kuliner semakin ketat, mendorong pentingnya kualitas produk, pelayanan, dan citra merek di mata pelanggan [1]. Salah satu restoran populer dikalangan masyarakat saat ini yang mengalami kemajuan pesat adalah merek Mie Gacoan, yang banyak diminati semua kalangan. Jaringan restoran mie pedas bermerek Mie Gacoan, yang kini menduduki peringkat teratas di Indonesia. Usaha ini adalah bagian dari PT Pesta Pora Abadi dan mulai beroperasi sejak awal tahun 2016 [2]. Seiring berjalannya waktu, Mie Gacoan terus meningkat dan membuka banyak cabang di berbagai wilayah, termasuk di wilayah Sulawesi Barat, baru-baru ini telah dibuka cabang tepatnya di kota Mamuju untuk pertama kalinya membuka Mie Gacoan. Kehadiran merek nasional seperti Mie Gacoan di daerah ini tentu menimbulkan respon yang beragam dari masyarakat lokal, baik dalam bentuk antusiasme maupun kritik melalui ulasan pelanggan. Meskipun demikian, sampai saat ini belum ada penelitian yang membahas persepsi awal konsumen di wilayah Sulawesi Barat terhadap kehadiran Mie Gacoan. Padahal, pemahaman terhadap tanggapan konsumen lokal sangat penting sebagai bahan evaluasi kualitas layanan dan menyesuaikan strategi bisnis berdasarkan kebutuhan dan persepsi pelanggan, khususnya di daerah yang baru dijangkau. Oleh karena itu, diperlukan pendekatan ilmiah untuk menganalisis ulasan dan opini pelanggan secara sistematis, agar diketahui bagaimana penerimaan masyarakat terhadap restoran tersebut, salah satu cara memanfaatkan ulasan pelanggan di platform online adalah dengan melakukan analisis sentimen.

Analisis sentimen dimaksudkan untuk memahami kecenderungan pendapat dari sekelompok masyarakat atau individu, biasanya dalam bentuk tulisan atau komentar. Analisis sentimen, mampu memperoleh informasi penting dari data yang tidak terstruktur, di mana informasi dihasilkan berupa pendapat yang dikategorikan ke dalam sentimen positif atau negatif berdasarkan emosi yang terkandung dalam pendapat tersebut [3]. Penerapan analisis sentimen memerlukan pendekatan khusus melalui pemilihan metode yang paling sesuai. Dalam penelitian ini, digunakan metode klasifikasi, yaitu teknik pengelompokan data berdasarkan kelas yang telah ditentukan sebelumnya. Algoritma klasifikasi yang diterapkan dalam penelitian ini ialah *Multinomial Logistic Regression*, yang merupakan suatu bentuk pengembangan dari *Logistic Regression* untuk menyelesaikan masalah klasifikasi dengan lebih dari dua kategori. Secara umum, *Logistic Regression* merupakan metode pembelajaran terawasi (*supervised learning*) yang mampu menyelesaikan permasalahan regresi maupun klasifikasi [4] *Logistic Regression* adalah salah satu cara regresi sederhana yang dimaksudkan untuk klasifikasi. Metode ini memodelkan probabilitas suatu kelas output guna memahami hubungan antara beberapa variabel, di mana variabel respon bersifat kategorikal, sedangkan variabel penjelas dapat berupa kategorikal maupun kontinu. [5] Konsep dasar tersebut menjadi landasan bagi penggunaan *Multinomial Logistic Regression* dalam penelitian ini.

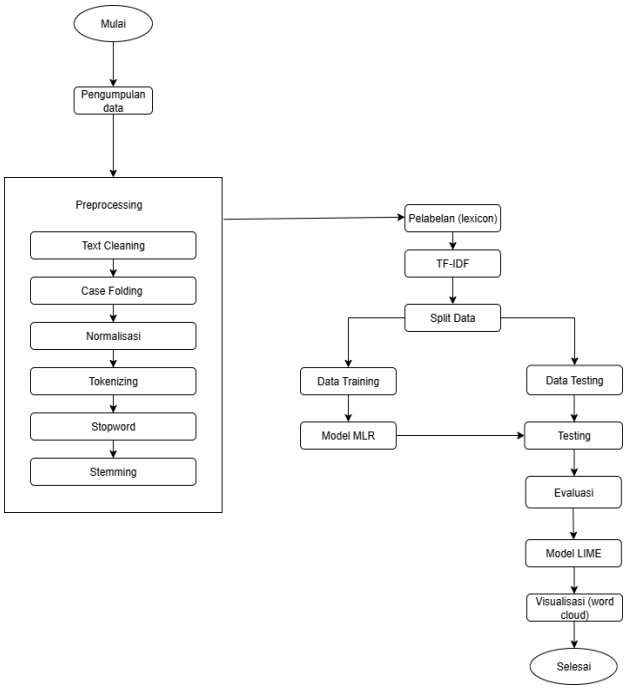
Berdasarkan penelitian sebelumnya, sejumlah studi menunjukkan bahwa *Multinomial Logistic Regression* mampu memberikan performa klasifikasi yang baik. Penelitian sebelumnya dilakukan oleh Assaidi dan Amin [6] Hasil penelitian menunjukkan algoritma *Logistic Regression* dapat menghasilkan performa klasifikasi yang baik dengan akurasi 78,57%, presisi 76,92%, recall 83,3%, dan f1-score 80%. Penelitian oleh Hidayati, Arief Senja, dan Mochamad Alfan [7] menggunakan metode *Multinomial Logistic Regression* untuk mengklasifikasikan ke dalam tiga kategori sentimen, yakni positif, negatif, serta netral dengan akurasi mencapai 86%. Penelitian sebelumnya juga dilakukan oleh Fathur Rizal, Andi Wijaya, dan Fuadz Hasyim [8] penelitian ini berhasil membangun model klasifikasi menggunakan *Logistic Regression* dengan akurasi tertinggi sebesar 83%. Penelitian selanjutnya juga dilakukan oleh Rahmawati dan Fitriani [9] hasil penelitian menunjukkan bahwa algoritma *Logistic Regression* menghasilkan performa klasifikasi

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI : 10.62411/tc.v25i1.15532

yang paling optimal dibandingkan algoritma Naïve Bayes dan Random Forest, dengan akurasi sebesar 82%, presisi 82%, recall 78%, dan f1-score 80%. Berdasarkan temuan-temuan tersebut, *Multinomial Logistic Regression* dengan pembobotan TF-IDF banyak digunakan dalam analisis sentimen, serta mulai dipadukan dengan *Explainable AI* seperti *Local Interpretable Model-agnostic Explanations* (LIME). Namun, umumnya penelitian tersebut masih memanfaatkan satu sumber data ulasan dan belum mengaitkan hasil interpretasi model dengan konteks wilayah penelitian. Penelitian ini menghadirkan kebaruan dengan mengintegrasikan ulasan pelanggan dari Google Maps dan data kuesioner yang dianalisis menggunakan *Multinomial Logistic Regression* berbasis TF-IDF serta dilengkapi pendekatan *Local Interpretable Model-agnostic Explanations* (LIME). Pendekatan ini digunakan untuk mengklasifikasikan sentimen sekaligus menjelaskan faktor kata yang berpengaruh pada konteks restoran Mie Gacoan di Kota Mamuju. Secara umum, penelitian ini memberikan tiga kontribusi utama, yaitu integrasi data ulasan dari Google Maps dan kuesioner untuk memperoleh gambaran sentimen yang lebih komprehensif pada konteks restoran lokal, penerapan *Multinomial Logistic Regression* berbasis pembobotan TF-IDF untuk mengklasifikasikan sentimen ke dalam kategori positif, negatif, dan netral, serta pemanfaatan pendekatan *Explainable AI* melalui metode *Local Interpretable Model-agnostic Explanations* (LIME) guna mengidentifikasi kata-kata yang berkontribusi signifikan terhadap hasil klasifikasi dan meningkatkan interpretabilitas model pada konteks wilayah Kota Mamuju. Berdasarkan latar belakang yang telah diuraikan, penelitian ini bertujuan untuk melakukan analisis sentimen pelanggan restoran Mie Gacoan di Kota Mamuju menggunakan algoritma *Multinomial Logistic Regression*, serta memanfaatkan pendekatan *Explainable AI* melalui metode LIME untuk memberikan penjelasan terhadap hasil klasifikasi sentimen dan mengidentifikasi faktor-faktor utama yang memengaruhi persepsi pelanggan berdasarkan ulasan restoran.

2. METODE PENELITIAN

Metode penelitian ini meliputi beberapa tahapan mulai dari pengumpulan data, *preprocessing,* pelabelan, pembobotan kata dengan TF-IDF, split data, klasifikasi, evaluasi, penerapan model LIME dan visualisasi. Gambar 1 adalah alur penelitian yang telah dilakukan.



Gambar 1. Tahapan penelitian

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI : 10.62411/tc.v25i1.15532

2.1. Pengumpulan data

Data penelitian diperoleh dari dua sumber utama, yaitu ulasan pelanggan dari Google Maps dan data kuesioner pelanggan. Data ulasan Google Maps dikumpulkan menggunakan teknik web scraping dengan memanfaatkan layanan SerpApi pada halaman restoran yang berlokasi di wilayah Kota Mamuju. Data kuesioner dikumpulkan secara daring melalui Google Form, di mana setiap responden memberikan satu ulasan dalam bentuk teks yang selanjutnya diperlakukan sebagai satu dokumen dalam proses analisis. Pengambilan sampel dilakukan dengan teknik convenience sampling, yaitu responden yang bersedia berpartisipasi berdasarkan pengalaman kunjungan mereka ke restoran. Kuesioner disusun dalam bentuk pertanyaan terbuka agar responden dapat menyampaikan pendapat dan pengalaman mereka secara bebas terkait layanan dan kualitas restoran. Seluruh jawaban yang diperoleh kemudian digunakan sebagai data ulasan pelanggan dan digabungkan dengan data ulasan dari Google Maps ke dalam satu berkas CSV untuk selanjutnya diproses pada tahap analisis.

2.2 *Preprocessing*

Bagian ini merupakan langkah dalam proses pembersihan data yang bertujuan mengubah data ke dalam format yang lebih terstruktur sehingga proses data mining dapat dilakukan secara optimal [10] Beberapa tahapan dilakukan dalam fase ini untuk memastikan data siap digunakan adapun langkah-langkah prosesnya sebagai berikut:

1) *Text cleaning* proses untuk menghapus elemen yang tidak dibutuhkan dari data teks, seperti tautan (URL), emoji, simbol, dan tanda baca.
2) *Case Folding* seluruh huruf dalam teks diubah menjadi huruf kecil (lowercase).
3) *Normalisasi* mengubah teks yang mengandung singkatan atau kata-kata yang tidak baku menjadi ejaan yang benar.
4) *Tokenizing* memecah teks menjadi token.
5) *Stopword* menghapus teks umum yang tidak memiliki makna penting
6) *Stemming* memotong atau menghilangkan imbuhan pada teks agar diperoleh bentuk dasar.

2.3 Pelabelan Data

Pada tahapan ini digunakan pelabelan *lexicon based* yang berisi daftar kata yang telah dikategorikan sebagai positif, negatif dan juga netral. Kata-kata dalam teks akan dibandingkan dengan *lexicon based* tersebut untuk menentukan sentimen [11]. Dalam penelitian ini, kamus leksikon oleh Massacex [12] digunakan sebagai acuan untuk menentukan polaritas setiap kata pada teks ulasan yang telah melalui tahap *preprocessing*. Penentuan kategori sentimen dilakukan dengan menjumlahkan bobot setiap kata dalam teks ulasan. Setiap kata yang ditemukan dalam kamus leksikon memiliki bobot sentimen tertentu, sedangkan kata yang tidak terdapat dalam kamus diberi nilai nol. Skor total sentimen suatu ulasan diperoleh dari penjumlahan seluruh bobot kata, kemudian kategori sentimen ditentukan berdasarkan threshold berbasis nilai nol. Dalam penelitian ini, tidak diterapkan mekanisme khusus untuk menangani struktur negasi secara kontekstual, seperti frasa “tidak enak” atau “kurang baik”. Kata-kata diproses secara individual berdasarkan bobot yang tersedia dalam kamus leksikon, sehingga pengaruh negasi mengikuti bobot masing-masing kata yang terdapat dalam teks.

$$if \sum_k wk > 0 \rightarrow Positif \quad (1)$$

$$if \sum_k wk < 0 \rightarrow Negatif \quad (2)$$

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI : 10.62411/tc.v25i1.15532

$$if \sum_k wk = 0 \rightarrow Netral \quad (3)$$

Keterangan:
Wk : Bobot sentimen kata ke-k berdasarkan kamus leksikon.
$\sum_k wk$: Skor total sentimen ulasan, yaitu hasil penjumlahan bobot seluruh kata dalam teks.

2.4 *Term Frequency-Inverse Document Frequency* (TF-IDF)

Melalui metode ini, teks direpresentasikan dalam bentuk numerik sehingga dapat diproses oleh algoritma pembelajaran mesin. Pendekatan TF-IDF digunakan untuk menilai tingkat kepentingan suatu kata dalam sebuah dokumen dengan mempertimbangkan distribusinya pada keseluruhan kumpulan dokumen. *Term Frequency* (TF) merefleksikan frekuensi kemunculan kata pada dokumen tertentu, sedangkan *Inverse Document Frequency* (IDF) menunjukkan sebaran kata tersebut dalam seluruh dokumen. Kombinasi kedua komponen tersebut menghasilkan bobot kata yang mencerminkan tingkat relevansinya, di mana kata yang sering muncul pada satu dokumen namun jarang ditemukan pada dokumen lain akan memperoleh nilai bobot yang lebih tinggi. [13]

1) *Term Frequency (TF)*
Menunjukkan frekuensi kemunculan kata t dalam dokumen d.

$$TF(t,d) = \frac{ft,d}{\sum_k fk,d} \quad (4)$$

Keterangan:
ft,d : Jumlah kemunculan kata t di dokumen d
$\sum_k fk,d$: Jumlah keseluruhan kata yang terdapat dalam dokumen d

2) *Inverse Document Frequency (IDF)*
Menunjukkan tingkat kepentingan suatu kata t dalam keseluruhan kumpulan dokumen.

$$IDF(t) = log\left(\frac{N}{df_t}\right) \quad (5)$$

Keterangan:
N : Jumlah keseluruhan dokumen
df_t : Jumlah dokumen yang memuat kata

3) TF-IDF

$$TF - IDF(t,d) = TF(t,d) \times IDF(t) \quad (6)$$

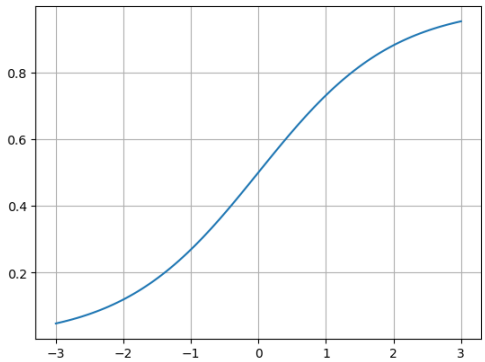
2.5 Split Data

Pada penelitian ini, dataset yang telah melalui proses *preprocessing* serta pemberian bobot fitur menggunakan TF-IDF selanjutnya dipisahkan ke dalam dua bagian, yaitu data latih sebesar 80% dan data uji sebesar 20%.

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI : 10.62411/tc.v25i1.15532

2.6 Klasifikasi *Multinomial Logistic Regression*

Multinomial Logistic Regression merupakan salah satu metode regresi sederhana yang banyak digunakan dalam permasalahan klasifikasi, yang mengonversi nilai masukan menjadi probabilitas.



Gambar 2. Kurva

Kurva yang terlihat pada Gambar 2 menggambarkan fungsi sigmoid pada *Logistic Regression* biner sebagai dasar konseptual, yang secara matematis dinyatakan sebagai:

$$\sigma\,(z) \;=\; \frac{1}{1 \;+\; e^{-z\prime}} \;\; z \;=\; w.x \;+\; b \qquad (7)$$

Keterangan:

$\sigma\,(z)$	: Fungsi sigmoid
z	: Kombinasi linier fitur ($z = w.x + b$)
w	: Vektor bobot model
x	: Vektor fitur teks
b	: Bias model
e	: Konstanta eksponensial

z merupakan hasil dari kombinasi liniar bobot dan fitur yang diberikan. Fungsi sigmoid mengkonversi nilai dari kombinasi liniar tersebut menjadi rentang probabilitas antara 0 dan 1, sehingga efektif untuk menggambarkan kemungkinan suatu kelas.

Konsep dari fungsi logistik ini menjadi pondasi bagi pengembangan *Multinomial Logistic Regression* (MLR) yang digunakan untuk memprediksi variabel dependen dengan lebih dari dua kategori. *Multinomial Logistic Regression*, yang juga dikenal sebagai Regresi *Softmax*, merupakan metode klasifikasi yang memperluas konsep *Logistic Regression* ke dalam kasus dengan lebih dari dua kelas. Metode ini digunakan untuk memperkirakan probabilitas dari berbagai hasil kategori yang berbeda berdasarkan sejumlah variabel independen [14]. Dalam penelitian ini, klasifikasi sentimen dilakukan menggunakan *Multinomial Logistic Regression.*

$$P\big(y = c \,\big|x\,;\; \theta_{1,}\, \theta_{2},,,,\theta_{k}\,\big) = \frac{exp\big(\theta_{c}^{T} x\big)}{\sum_{j=1}^{k} exp\big(\theta_{j}^{T} x\big)} \qquad (8)$$

Keterangan:

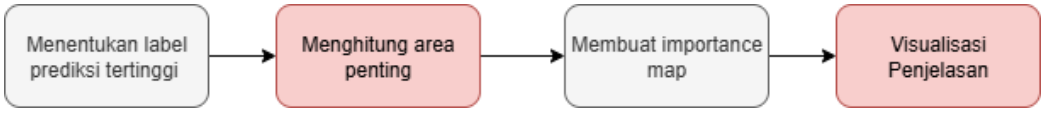
e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI : 10.62411/tc.v25i1.15532

$p(y = c \mid x; \theta)$: Probabilitas bahwa input *x* termasuk ke dalam kelas *c*
θ_c : Parameter (vektor bobot) untuk kelas c
$(\theta_c^T x)$: Hasil dot product antara parameter kelas c dan input x
Penyebut (denominator) adalah penjumlahan semua eksponensial dari dot product antara parameter setiap kelas dan input memastikan bahwa total probabilitas = 1

2.7 Evaluasi

Pada bagian ini digunakan *confusion matrix* sebagai evaluasi performa yang menunjukkan jumlah prediksi yang benar maupun salah, yang dapat digunakan untuk klasifikasi dua atau lebih kelas. Ukuran evaluasi seperti akurasi, presisi, recall, dan f1-score digunakan untuk menilai tingkat akurasi hasil prediksi dalam penelitian. [15]

2.8 *Local Interpretable Model-agnostic Explanations* (LIME)



Gambar 3. Tahapan LIME

Local Interpretable Model-agnostic Explanations (LIME) merupakan metode untuk menginterpretasikan model yang digunakan untuk memberikan penjelasan lokal terhadap hasil perkiraan dari model yang bersifat black box, tanpa bergantung pada jenis algoritma yang digunakan (model-agnostic). Teknik ini bekerja secara post-hoc, yaitu memberikan penjelasan setelah model selesai dibangun.

Dalam penerapannya pada data berbasis teks, LIME menggunakan kata sebagai fitur utama untuk menyoroti elemen-elemen yang paling berpengaruh terhadap keputusan prediksi model. Pada penelitian ini, LIME digunakan untuk menjelaskan alasan model *Multinomial Logistic Regression* mengklasifikasikan suatu teks ke dalam kategori sentimen positif, negatif, dan netral [16].

$$(x) = argmin_{g \in G} \{L(f, g, \pi) + \Omega\,(g)\} \quad (9)$$

Keterangan:
X : Sebuah model yang diprediksi Sebuah contoh dari ruang data yang diinginkan penjelasannya terhadap nilai target yang diprediksi oleh model.
L $(f, g, \pi x)$: Fungsi fidelitas yang digunakan untuk mengukur tingkat kesesuaian model penjelas (g) terhadap model asli (f) pada area lokal yang\ ditentukan oleh bobot kedekatan (πx).
$\Omega(g)$: Mengukur kompleksitas model dari penjelas (g)
f : Model kotak hitam yang akan dijelaskan
g : Penjelas
G : Total set model yang dapat ditafsirkan
πx : Ukuran kedekatan

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI : 10.62411/tc.v25i1.15532

3. HASIL DAN PEMBAHASAN

3.1 Pengumpulan Data

Berdasarkan proses pengumpulan data, diperoleh sebanyak 1651 ulasan dari Google Maps dan 356 data kuesioner pelanggan. Seluruh data tersebut kemudian digabungkan ke dalam satu file CSV, sehingga diperoleh 2007 data dan melalui tahap pembersihan serta *preprocessing* hingga diperoleh 773 sebagai dataset akhir yang siap diproses untuk analisis selanjutnya.

Tabel 1. Jumlah Data

No	Tahapan	Jumlah Data
1	Data awal (gabungan dua sumber)	2007
2	Data kosong dihapus	926
3	Data simbol/emotikon dihapus	2
4	Data duplikat dihapus	295
5	Dataset setelah pembersihan awal	784
6	Data tanpa token setelah *preprocessing* dihapus	11
7	Dataset akhir untuk analisis	**773**

3.2 *Preprocessing*

Preprocessing teks bertujuan untuk membuat data teks lebih terstruktur dengan menghilangkan teks yang tidak penting, sehingga mempermudah proses selanjutnya.

3.2.1 Cleaning

Proses ini membersihkan data teks dari berbagai elemen seperti tautan (URL), emoji, simbol, dan tanda baca yang tidak diperlukan atau dapat mengganggu proses analisis. Tahap ini bertujuan agar data ulasan yang diperoleh, baik dari Google Maps maupun kuesioner, menjadi lebih terstruktur dan mudah diproses. Pada tabel 2 merupakan hasil dari data cleaning yang sudah dibersihkan di mana pada opini pertama terdapat tanda seru (!) juga emoji, pada opini ke dua terdapat tanda petik koma (“.) koma atas bawah (‘,), tanda titik (.) dan terdapat angka yang telah dihapus, serta pada opini ketiga terdapat juga emoji yang telah dibersihkan.

Tabel 2. Tahap data *cleaning*

No	Sebelum *Cleaning*	Setelah *Cleaning*
1	Udang keju nya enak sekali!!! pokoknya the best deh, kalo pulang sekolah recommended banget singgah makan disini 😍□□	Udang keju nya enak sekali pokoknya the best deh kalo pulang sekolah recommended banget singgah makan disini

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI : 10.62411/tc.v25i1.15532

2	kok staf nya ngak datang". udah 3 jam loh nunggunya. kita kemari bukan untuk duduk kak', kita mau makan. sekali lagi kita ke gacoan mau MAKAN BUKAN MAU DUDUK	kok staf nya ngak datang udah jam loh nunggunya kita kemari bukan untuk duduk kak kita mau makan sekali lagi kita ke gacoan mau MAKAN BUKAN MAU DUDUK
3	Suasana nya sihh senyaman itu 🙏🙏	Suasana nya sihh senyaman itu

3.2.2 *Case Folding*

Pada bagian ini seluruh huruf dalam teks diubah menjadi huruf kecil (lowercase). Langkah tersebut dilakukan untuk menyeragamkan format penulisan kata sehingga sistem tidak membedakan kata yang sama hanya berdasarkan huruf yang berbeda, besar atau kecil, seperti kata “Enak” dan “enak”.

Tabel 3. Tahap data *Case folding*

No	Sebelum *Case folding*	Setelah *Case folding*
1	Udang keju nya enak sekali pokoknya the best deh kalo pulang sekolah recommended banget singgah makan disini	udang keju nya enak sekali pokoknya the best deh kalo pulang sekolah recommended banget singgah makan disini
2	kok staf nya ngak datang udah jam loh nunggunya kita kemari bukan untuk duduk kak kita mau makan sekali lagi kita ke gacoan mau MAKAN BUKAN MAU DUDUK	kok staf nya ngak datang udah jam loh nunggunya kita kemari bukan untuk duduk kak kita mau makan sekali lagi kita ke gacoan mau makan bukan mau duduk
3	Suasana nya sihh senyaman itu	suasana nya sihh senyaman itu

Pada tabel 3 merupakan hasil dari *case folding* yang di mana semua opini yang memiliki huruf kapital diubah menjadi huruf kecil agar memiliki keseragaman yang sama dalam format penulisannya.

3.2.3 *Normalization*

Pada bagian ini proses mengganti teks tidak baku atau singkatan dalam ulasan menjadi sesuai dengan ejaan yang benar. Contohnya, kata “gk” diubah menjadi “tidak”, tujuannya agar menjadi lebih rapi dan seragam.

Tabel 4. Tahap data *Normalization*

No	Sebelum *Normalization*	Setelah *Normalization*
1	udang keju nya enak sekali pokoknya the best deh kalo pulang sekolah recommended banget singgah makan disini	udang keju ya enak sekali pokoknya the best deh kalau pulang sekolah recommended banget singgah makan disini
2	kok staf nya ngak datang udah jam loh nunggunya kita kemari bukan untuk duduk kak kita mau	kok staf ya tidak datang sudah jam loh nunggunya kita kemari bukan untuk duduk kak kita mau makan sekali lagi

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI : 10.62411/tc.v25i1.15532

	makan sekali lagi kita ke gacoan mau makan bukan mau duduk	kita ke gacoan mau makan bukan mau duduk
3	suasana nya sihh senyaman itu	suasana ya sih senyaman itu

Pada tabel 4 merupakan hasil normalisasi yang di mana pada opini ke dua kata (ngak) dinormalkan menjadi (tidak) dan juga pada opini ke tiga kata (nya) dinormalkan menjadi (ya). Begitu pun pada kata-kata lain yang mengalami perubahan serupa.

3.2.4 *Tokenizing*

Pada Bagian tahapan ini bertujuan untuk memisahkan teks ke dalam unit-token yang umumnya berupa kata, frasa, atau bahkan karakter, tergantung pada kebutuhan analisis. contoh kalimat “Makanannya enak dan murah” akan dipecah menjadi token: ["makanannya", "enak", "dan", "murah"].

Tabel 5. Tahap data *tokenizing*

No	Sebelum *Tokenizing*	Sesudah *Tokenizing*
1	udang keju ya enak sekali pokoknya the best deh kalau pulang sekolah recommended banget singgah makan disini	['udang', 'keju', 'ya', 'enak', 'sekali', 'pokoknya', 'the', 'best', 'deh', 'kalau', 'pulang', 'sekolah', 'recommended', 'banget', 'singgah', 'makan', 'disini']
2	kok staf ya tidak datang sudah jam loh nunggunya kita kemari bukan untuk duduk kak kita mau makan sekali lagi kita ke gacoan mau makan bukan mau duduk	['kok', 'staf', 'ya', 'tidak', 'datang', 'sudah', 'jam', 'loh', 'nunggunya', 'kita', 'kemari', 'bukan', 'untuk', 'duduk', 'kak', 'kita', 'mau', 'makan', 'sekali', 'lagi', 'kita', 'ke', 'gacoan', 'mau', 'makan', 'bukan', 'mau', 'duduk']
3	suasana ya sih senyaman itu	['suasana', 'ya', 'sih', 'senyaman', 'itu']

Pada tabel 5 merupakan hasil proses dari tokenizing yang di mana kata-katanya dipecah menjadi bagian-bagian kecil contohnya pada opini pertama ['udang', 'keju', 'ya', 'enak', 'sekali', 'pokoknya', 'the', 'best', 'deh', 'kalau', 'pulang', 'sekolah', 'recommended', 'banget', 'singgah', 'makan', 'disini']. Begitu pun dengan kata-kata lainnya semuanya sama.

3.2.5 *Stopword*

Merupakan tahapan penyaringan terhadap kata-kata yang bersifat umum, seperti “yang”, “dan”, “di”, serta “untuk”, yang memiliki frekuensi kemunculan tinggi dalam teks namun tidak berperan signifikan dalam pembentukan makna sentimen

Tabel 6. Tahap data *stopword*

No	Sebelum *Stopword*	Setelah *Stopword*
1	['udang', 'keju', 'ya', 'enak', 'sekali', 'pokoknya', 'the', 'best', 'deh', 'kalau', 'pulang', 'sekolah', 'recommended', 'banget', 'singgah', 'makan', 'disini']	['udang', 'keju', 'enak', 'pokoknya', 'best', 'pulang', 'sekolah', 'recommended', 'banget', 'singgah', 'makan']
2	['kok', 'staf', 'ya', 'tidak', 'datang', 'sudah', 'jam', 'loh', 'nunggunya', 'kita', 'kemari', 'bukan', 'untuk', 'duduk', 'kak', 'kita', 'mau', 'makan', 'sekali', 'lagi', 'kita', 'ke',	['staf', 'jam', 'nunggunya', 'kemari', 'duduk', 'kak', 'makan', 'gacoan', 'makan', 'duduk']

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI : 10.62411/tc.v25i1.15532

	'gacoan', 'mau', 'makan', 'bukan', 'mau', 'duduk']	
3	['suasana', 'ya', 'sih', 'senyaman', 'itu']	['suasana', 'senyaman']

Pada tabel 6 merupakan hasil *Stopword* yang telah dibersihkan contohnya pada opini pertama kata "ya", "sekali", "the", "deh", "kalau", "disini". Kata tersebut dihapus karena tidak memiliki kontribusi penting dalam analisis, begitu pun dengan opini lainnya semua kata yang memiliki makna yang tidak penting akan dihapus.

3.2.6 Stemming

Bagian ini merupakan proses terakhir, yaitu stemming yang bertujuan untuk mengurangi atau menghilangkan imbuhan dari kata sehingga menjadi bentuk dasar (akar kata). Dalam ulasan pengguna, sering ditemukan variasi kata-kata yang bentuknya berbeda tetapi memiliki arti yang sama, seperti "bermain", "dimainkan", atau "memainkan". Ketiganya akan dikembalikan menjadi satu kata dasar, yaitu "main".

Tabel 7. Tahap data *stemming*

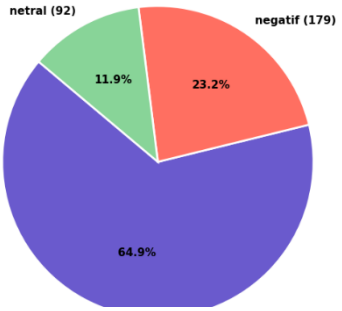
No	Sebelum *Stemming*	Setelah *Stemming*
1	['udang', 'keju', 'enak', 'pokoknya', 'best', 'pulang', 'sekolah', 'recommended', 'banget', 'singgah', 'makan']	udang keju enak pokok best pulang sekolah recommended banget singgah makan
2	['staf', 'jam', 'nunggunya', 'kemari', 'duduk', 'kak', 'makan', 'gacoan', 'makan', 'duduk']	staf jam nunggu kemari duduk kak makan gaco makan duduk
3	['suasana', 'senyaman']	suasana nyaman

Pada tabel 7 merupakan hasil stemming yang di mana setiap kata akan diubah ke bentuk dasar contohnya pada opini pertama kata (pokoknya) diubah menjadi (pokok), pada opini ke 2 kata (nunggunya) diubah menjadi (nunggu) begitu pun dengan opini lainnya semuanya diubah menjadi kata dasar.

3.3 Pelabelan Data

Setelah selesai tahap pembersihan data atau preprocessing didapatkan data bersih sebanyak 773 data dari data sebelumnya sebanyak 2007 data, tahap selanjutnya yaitu pelabelan data dengan menggunakan pendekatan *lexicon-based,* digunakan untuk mengklasifikasikan setiap ulasan pelanggan ke dalam tiga kategori sentimen, yaitu positif, negatif, dan netral.

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI : 10.62411/tc.v25i1.15532

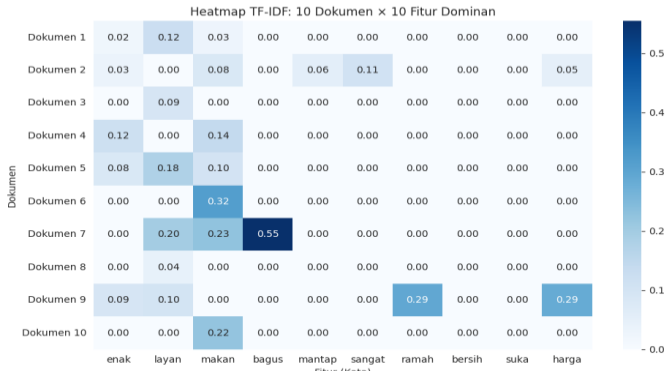


Gambar 4. Pelabelan

Pada Gambar 4 di atas merupakan tampilan dari hasil pelabelan *lexicon based* Berdasarkan hasil pelabelan, diperoleh distribusi data sentimen positif sebanyak 502 data 64,9%, menunjukkan bahwa sebagian besar pelanggan memberikan tanggapan yang baik terhadap layanan dan produk restoran, sentimen negatif sebanyak 179 data 23,2% terlihat adanya sejumlah ulasan yang berisi keluhan atau ketidakpuasan pelanggan, dan sentimen netral sebanyak 92 data 11,9% berisi ulasan yang bersifat informatif tanpa menunjukkan emosi positif maupun negatif secara jelas. Hasil distribusi ini menunjukkan bahwa banyak pelanggan memiliki persepsi positif terhadap restoran, namun masih terdapat proporsi ulasan negatif yang dapat menjadi perhatian bagi pihak restoran dalam meningkatkan kualitas layanan.

3.4 Pembobotan kata *TF-IDF* dan Pembagian data

Tahap berikutnya setelah pelabelan adalah melakukan proses penentuan bobot kata, metode *Term Frequency–Inverse Document Frequency* (TF-IDF) diterapkan untuk merepresentasikan data teks ke dalam data numerik yang dapat diklasifikasikan. Kata dengan frekuensi kemunculan yang lebih tinggi akan memperoleh bobot yang lebih besar dibandingkan kata lainnya, menurut hasil pembobotan. Setelah proses pembobotan selesai, data dibagi menjadi 80:20, dengan 80 persen dialokasikan untuk data pelatihan dan 20 persen dialokasikan untuk data testing, pembagian ini bertujuan untuk memperoleh akurasi dari kinerja model yang dibangun.



Gambar 5. TF-IDF

3.5 Klasifikasi *Multinomial Logistic Regression*

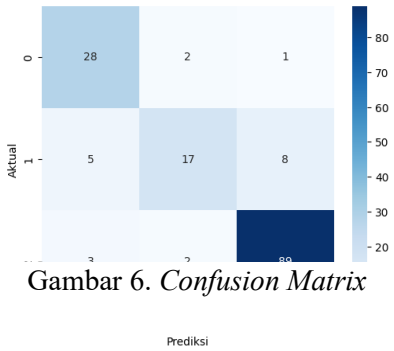
Setelah melalui tahap *preprocessing*, pelabelan, pembobotan kata, dan pembagian data selanjutnya Model *Multinomial Logistic Regression* dilatih menggunakan data hasil pembobotan TF-IDF untuk mempelajari pola hubungan antara kata dan jenis sentimen. Selama proses pelatihan, parameter model (θ) disesuaikan agar menghasilkan prediksi paling akurat melalui proses optimasi dan validasi sederhana.

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI : 10.62411/tc.v25i1.15532

Model ini kemudian digunakan untuk memprediksi ke dalam tiga kategori sentimen, meliputi positif, negatif, dan netral pada data pengujian.

3.6 Evaluasi

Tahap selanjutnya adalah evaluasi yang dilakukan melalui analisis *confusion matrix* guna mengevaluasi performa model yang diterapkan, *Multinomial Logistic Regression* bekerja dalam mengklasifikasikan tiga kelas sentimen, yaitu positif, negatif, dan netral.



Gambar 6. *Confusion Matrix*

Dari 31 data negatif, model memprediksi 28 data dengan benar sebagai negatif, 2 data benar dianggap netral, dan 1 data dengan benar dianggap positif, seperti yang ditunjukkan pada Gambar 6 *confusion matrix*, dari 30 data netral, model memprediksi 17 data dengan benar sebagai netral, tetapi 5 data dengan benar dianggap negatif, dan 8 data dengan benar dianggap positif. Sementara itu, dari data positif, model mengklasifikasikan 89 data dengan benar, 3 data salah diklasifikasikan negatif dan juga 2 data salah diklasifikasikan sebagai netral.

Tabel 8. Hasil Evaluasi

Model Multinomial Logistic Regression	*Accuracy*	*Precision*	*Recall*	*F1-score*
	86,5%	86,3%	86,5%	85,8%

Berdasarkan tabel 8 hasil evaluasi, model *Multinomial Logistic Regression* memiliki akurasi 86,5%, presisi 86,3%, recall 86,5%, dan F1-score 85,8%. Nilai menunjukkan bahwa pendekatan yang diterapkan memiliki kemampuan yang baik dalam mengklasifikasikan sentimen pelanggan secara akurat dan konsisten, sehingga layak digunakan sebagai alat analisis opini pelanggan restoran.

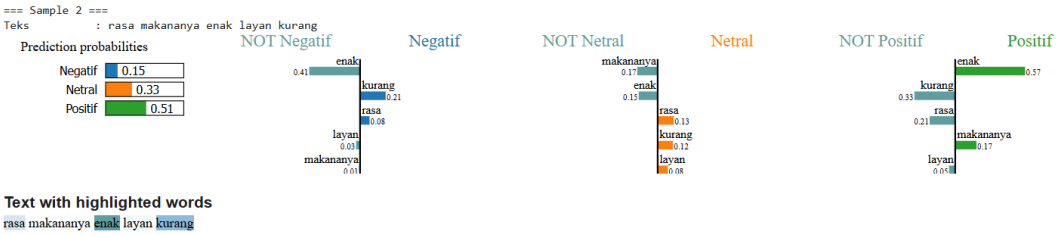
3.7 *Local Interpretable Model-agnostic Explanations* (LIME)

Setelah proses evaluasi kinerja model dilakukan menggunakan *confusion matrix* yang mencakup nilai akurasi, presisi, recall, dan F1-score, tahapan berikutnya adalah menginterpretasikan hasil klasifikasi dengan menerapkan metode *Local Interpretable Model-agnostic Explanations* (LIME). Metode LIME menghasilkan visualisasi interpretatif sekaligus memberikan informasi mengenai kontribusi setiap kata terhadap keputusan klasifikasi sentimen yang dihasilkan oleh algoritma. Informasi tersebut digunakan untuk mengetahui bagian teks yang memiliki pengaruh paling besar dalam penentuan kategori sentimen, baik positif, negatif, maupun netral.

Proses implementasi LIME pada penelitian ini diawali dengan mengimpor pustaka lime.lime_text menggunakan perintah from lime.lime_text import *LimeTextExplainer*. Selanjutnya, ditetapkan tiga kelas sentimen, yaitu positif, negatif, netral yang disimpan dalam variabel class_names. Proses prediksi dilakukan dengan

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI : 10.62411/tc.v25i1.15532

memanfaatkan model klasifikasi yang sebelumnya telah dilatih, melalui fungsi predict_fn = lambda x: mlr_model.predict_proba(vectorizer.transform(x)). Fungsi ini bertugas menghitung probabilitas dari setiap kelas berdasarkan representasi teks hasil transformasi TF-IDF.



Gambar 7. LIME

Hasil LIME yang terdapat pada gambar 7 terhadap teks "rasa makanannya enak layan kurang", model mengklasifikasikan teks tersebut sebagai sentimen positif dengan probabilitas 51% sementara sentimen netral sebesar 33% dan negatif sebesar 15%. Visualisasi LIME menunjukkan bahwa kata "enak" memberikan kontribusi paling besar terhadap kelas positif, yang ditunjukkan oleh bar berwarna hijau dengan nilai kontribusi tertinggi. Selain itu, kata "makanannya" juga turut memperkuat kecenderungan ke arah sentimen positif, meskipun dengan kontribusi kata yang lebih kecil. Di sisi lain, kata "rasa "kurang" dan "layan" memiliki kontribusi yang cenderung mengarah ke kelas netral dan negatif, Namun, pengaruh kata-kata tersebut tidak lebih dominan dibandingkan kontribusi kata "enak", LIME menjelaskan bahwa dominasi kata-kata bernada positif menjadi alasan utama model mengklasifikasikan teks ini sebagai sentimen positif.

Gambar 8. *WordCloud LIME*

Pada gambar 8 memperlihatkan visualisasi *WordCloud* yang dihasilkan dari interpretasi LIME terhadap 10 sampel data uji. *WordCloud* ini menampilkan kata-kata yang memiliki kontribusi paling besar terhadap hasil prediksi model *Multinomial Logistic Regression*. *WordCloud* ini merepresentasikan kata-kata yang memiliki kontribusi paling besar dalam memengaruhi keputusan prediksi sentimen oleh model, di mana ukuran kata menunjukkan tingkat pengaruhnya. Terlihat bahwa kata "layan", "recommended", dan "keren" mendominasi tampilan *WordCloud* dengan ukuran yang paling besar. Hal ini mengindikasikan bahwa kata-kata tersebut memiliki pengaruh yang signifikan, khususnya dalam mendorong model ke arah sentimen positif. Selain itu, kata "banget" juga muncul cukup dominan, yang menunjukkan adanya penekanan ekspresi positif dalam ulasan pelanggan. Di sisi lain, kata "kurang" juga tampak dengan ukuran relatif besar, yang mencerminkan keberadaan unsur keluhan atau penilaian negatif dalam beberapa ulasan. Namun, kemunculan kata negatif ini tidak lebih dominan dibandingkan

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI : 10.62411/tc.v25i1.15532

kata-kata bernada positif seperti “recommended” dan “keren”, sehingga secara keseluruhan model tetap lebih sensitif terhadap ekspresi positif. Kata lain seperti “tempat”, “rasa”, “mie”, “bos”, dan “pokok” turut muncul sebagai kata pendukung yang berkontribusi dalam pembentukan konteks ulasan, baik positif maupun netral. Pola ini konsisten dengan hasil analisis LIME sebelumnya, di mana kata-kata bernada positif cenderung memiliki bobot kontribusi yang lebih besar dibandingkan kata negatif atau netral. Kondisi tersebut menunjukkan bahwa model memiliki kecenderungan lebih peka terhadap ekspresi positif, sehingga pada beberapa kasus, ulasan yang mengandung campuran pujian dan kritik tetap dapat diklasifikasikan sebagai sentimen positif. Meskipun demikian, penggunaan LIME berhasil memberikan interpretasi visual yang jelas dan informatif terkait kata-kata yang paling berpengaruh dalam proses pengambilan keputusan model.

3.8 WordCloud

Gambar 9. *WordCloud* sentimen positif

Gambar 10. *WordCloud* sentimen negatif

Gambar 11. *WordCloud* sentimen netral

Pada Gambar 9 menunjukkan kata-kata dari opini positif seperti enak, bersih, mantap, nyaman, bagus, ramah, yang menunjukkan opini positif masyarakat tentang kualitas layanan, rasa di restoran mie gacoan, gambar 10 menunjukkan kata-kata negatif seperti tidak, lama, layan, makan, pesan yang menunjukkan opini negatif masyarakat terhadap kualitas layanan restoran, gambar 11 menunjukkan kata-kata netral seperti makan, bersih, bagus, keren yang menunjukkan opini netral terhadap kualitas layanan restoran.

4. KESIMPULAN

Penelitian ini menunjukkan bahwa algoritma *Multinomial Logistic Regression* mampu mengklasifikasikan sentimen pelanggan restoran dengan baik, dengan nilai akurasi sebesar 86,5%, presisi 86,3%, recall 86,5%, dan F1-score 85,8%. Penerapan metode *Explainable AI* menggunakan LIME memberikan penjelasan yang transparan terhadap hasil klasifikasi sentimen dengan menyoroti kata-kata yang memiliki pengaruh paling besar dalam menentukan keputusan model. Berdasarkan hasil interpretasi LIME, kata-kata seperti “enak”, “recommended”, dan “keren” berkontribusi kuat dalam mendorong klasifikasi sentimen positif, sementara kata-kata seperti “layan”, “lama”,

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI : 10.62411/tc.v25i1.15532

"kurang", dan "tunggu" cenderung mengarah pada sentimen negatif. Selain itu, kata-kata seperti "makan", "mie", "tempat", dan "rasa" muncul sebagai kata pendukung yang membentuk ulasan yang ditemukan pada sentimen netral. Dengan demikian, penggunaan LIME membantu menjelaskan proses pengambilan keputusan model serta meningkatkan interpretabilitas hasil klasifikasi. Berdasarkan hasil tersebut, faktor utama yang mempengaruhi sentimen pelanggan adalah rasa makanan, rasa makanan yang enak membuat pelanggan suka makan di sana. Pelayanan restoran perlu ditingkatkan lamanya pelayanan kadang membuat pelanggan malas untuk menunggu lama, serta kenyamanan membuat pelanggan betah ada di restoran. Secara keseluruhan, kualitas rasa dan kecepatan pelayanan menjadi faktor paling dominan dalam membentuk persepsi pelanggan terhadap restoran. Hasil analisis sentimen menunjukkan bahwa sebagian besar ulasan pelanggan bersifat positif, yang menunjukkan penerimaan masyarakat yang baik terhadap kehadiran Mie Gacoan di wilayah Mamuju. Meskipun demikian, masih ditemukan sejumlah ulasan negatif yang berkaitan dengan aspek layanan, sehingga dapat menjadi bahan evaluasi bagi pihak restoran dalam meningkatkan kualitas pelayanan. Penelitian ini diharapkan dapat menjadi rujukan dalam pengembangan sistem analisis sentimen yang memiliki tingkat akurasi lebih baik, mudah diinterpretasikan, serta sesuai dengan kebutuhan pengguna di masa mendatang. Penelitian selanjutnya disarankan untuk melakukan perbandingan kinerja *Multinomial Logistic Regression* dengan algoritma klasifikasi lain, seperti *Support Vector Machine*, *Random Forest*, maupun pendekatan berbasis *deep learning*, guna memperoleh hasil klasifikasi sentimen yang lebih optimal. Selain itu, perlu dilakukan penambahan jumlah data serta penerapan skema pelabelan yang berbeda agar performa model dapat ditingkatkan. Penelitian selanjutnya juga dapat mengembangkan mekanisme penanganan struktur negasi dan intensifier melalui penerapan aturan linguistik tertentu, sehingga frasa dapat diinterpretasikan secara lebih akurat dalam proses analisis sentimen.

REFERENSI

[1] Y. A. Singgalen, "Analisis Sentimen Konsumen terhadap Food, Services, and Value di Restoran dan Rumah Makan Populer Kota Makassar Berdasarkan Rekomendasi Tripadvisor Menggunakan Metode CRISP-DM dan SERVQUAL," *Building of Informatics, Technology and Science (BITS)*, vol. 4, no. 4, Mar. 2023, doi: 10.47065/bits.v4i4.3231.

[2] G. A. Prasasti and P. Maisara, "Pengaruh Fasilitas, Harga Dan Cita Rasa Terhadap Kepuasan Konsumen Mie Gacoan Di Solo Raya," 2022, [Online]. Available: www.miegacoan.com

[3] B. Ramadhani and R. R. Suryono, "Komparasi Algoritma Naïve Bayes dan Logistic Regression Untuk Analisis Sentimen Metaverse," *Jurnal Media Informatika Budidarma*, vol. 8, no. 2, p. 714, Apr. 2024, doi: 10.30865/mib.v8i2.7458.

[4] F. Handayani *et al.*, "JEPIN (Jurnal Edukasi dan Penelitian Informatika) Komparasi Support Vector Machine, Logistic Regression Dan Artificial Neural Network dalam Prediksi Penyakit Jantung," 2021.

[5] N. Zatadinia and H. Dwi Bhakti, "Identifikasi Jamur Yang Dapat Dikonsumsi Dan Beracun Menggunakan Metode Logistic Regression," 2024.

[6] S. A. Assaidi and F. Amin, "Analisis Sentimen Evaluasi Pembelajaran Tatap Muka 100 Persen pada Pengguna Twitter menggunakan Metode Logistic Regression," 2022.

[7] A. R. Hidayati, A. S. Fitrani, M. A. Rosid, F. Sains, and D. Teknologi, "Analisa Sentimen Pemilu 2019 Pada Judul Berita Online Menggunakan Metode Logistic Regression," 2023.

[8] F. Rizal, A. Wijaya, and F. Hasyim, "Analisis Sentimen Masyarakat Indonesia Terhadap Aplikasi TikTok Menggunakan Algoritma Logistic Regression," *Journal homepage:*

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI : 10.62411/tc.v25i1.15532

Akiratech : Journal of Computer and Electrical Engineering, vol. 1, no. 2, 2024, [Online]. Available: https://journal.ajbnews.com/index.php/akiratech

[9] I. Rahmawati and T. R. Fitriani, "Analisis Sentimen Menggunakan Algoritma Logistic Regression Pada Penerbangan Lion Air berdasarkan Ulasan Pengguna Platform Online," 2023.

[10] H. Junianto, R. E. Saputro, B. A. Kusuma, D. Intan, and S. Saputra, "Comparison Of Logistic Regression And Random Forest In Sentiment Analysis Of Disdukcapil Application Reviews," *Jurnal Teknik Informatika (JUTIF)*, vol. 5, no. 6, 2024, doi: 10.52436/1.jutif.2024.5.6.1802.

[11] D. N. N. Septian Wulandari, "Analisis Sentimen Fitur Ulasan di Google Maps Pada Restoran Dapur Ingkung," 2024.

[12] Massacex, "GitHub - massacex/sentimen_analisis: analisis sentimen naive bayes lexicon based." Accessed: Feb. 14, 2026. [Online]. Available: https://github.com/massacex/sentimen_analisis

[13] A. Kartika Sari, Akhmad Irsyad, Dinda Nur Aini, Islamiyah, and Stephanie Elfriede Ginting, "Analisis Sentimen Twitter Menggunakan Machine Learning untuk Identifikasi Konten Negatif," *Adopsi Teknologi dan Sistem Informasi (ATASI)*, vol. 3, no. 1, pp. 64–73, Jun. 2024, doi: 10.30872/atasi.v3i1.1373.

[14] I. Qutab, K. I. Malik, and H. Arooj, "Sentiment Classification Using Multinomial Logistic Regression on Roman Urdu Text," *International Journal of Innovations in Science and Technology*, vol. 4, no. 2, pp. 323–335, Apr. 2022, doi: 10.33411/ijist/2022040204.

[15] A. V. T. R. Rihan Maulana, "Analisis Sentimen Ulasan Aplikasi Mypertamina Pada Google Play Store Menggunakan Algoritma Nbc," 2023.

[16] J. Khatib Sulaiman, N. Rizky Nuraeda, M. Liebenlito, and T. Edy Sutanto, "Explainable Sentiment Analysis pada Ulasan Aplikasi Shopee Menggunakan Local Interpretable Model-agnostic Explanations," *Indonesian Journal of Computer Science*, 2024.

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI : 10.62411/tc.v25i1.15532