

**PENERAPAN *LOGISTIC REGRESSION* UNTUK ANALISIS SENTIMEN PADA
ULASAN PRODUK *SKINCARE* DI *E-COMMERCE***

SKRIPSI

**Oleh :
JINAN NAURAH
NIM. 210605110136**



**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2025**

**PENERAPAN *LOGISTIC REGRESSION* UNTUK ANALISIS SENTIMEN
PADA ULASAN PRODUK *SKINCARE* DI *E-COMMERCE***

SKRIPSI

Diajukan kepada:
Universitas Islam Negeri Maulana Malik Ibrahim Malang
Untuk memenuhi Salah Satu Persyaratan dalam
Memperoleh Gelar Sarjana Komputer (S.Kom)

Oleh :
JINAN NAURAH
NIM. 210605110136

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2025**

HALAMAN PERSETUJUAN

PENERAPAN *LOGISTIC REGRESSION* UNTUK ANALISIS SENTIMEN PADA ULASAN PRODUK *SKINCARE* DI *E-COMMERCE*

SKRIPSI

Oleh :
JINAN NAURAH
NIM. 210605110136

Telah Diperiksa dan Disetujui untuk Diuji:
Tanggal: 28 November 2025

Pembimbing I,



Prof. Dr. Muhammad Faisal, M.T
NIP. 19740510 200501 1 007

Pembimbing II,



Dr. Ririen Kusumawati, S.Si, M.Kom
NIP. 19720309 200501 2 002

Mengetahui,
Ketua Program Studi Teknik Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang


Supriyono, M.Kom
NIP. 19841010 201903 1 012

HALAMAN PENGESAHAN

PENERAPAN *LOGISTIC REGRESSION* UNTUK ANALISIS SENTIMEN PADA ULASAN PRODUK *SKINCARE* DI *E-COMMERCE*





SKRIPSI

Oleh :
JINAN NAURAH
NIM. 210605110136

Telah Dipertahankan di Depan Dewan Penguji Skripsi
dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan
Untuk Memperoleh Gelar Sarjana Komputer (S.Kom)
Tanggal: 15 Desember 2025

Susunan Dewan Penguji

Ketua Penguji	: <u>Supriyono, M.Kom</u> NIP. 19841010 201903 1 012
Anggota Penguji I	: <u>Hani Nurhayati, M.T</u> NIP. 19780625 200801 2 006
Anggota Penguji II	: <u>Prof. Dr. Muhammad Faisal, M.T</u> NIP. 19740510 200501 1 007
Anggota Penguji III	: <u>Dr. Ririen Kusumawati, S.Si, M.Kom</u> NIP. 19720309 200501 2 002

()
()
()
()

Mengetahui dan Mengesahkan,
Ketua Program Studi Teknik Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang



Supriyono, M.Kom
NIP. 19841010 201903 1 012

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini:

Nama : Jinan Naurah
NIM : 210605110136
Fakultas / Program Studi : Sains dan Teknologi / Teknik Informatika
Judul Skripsi : Penerapan *Logistic Regression* Untuk Analisis Sentimen Pada Ulasan Produk *Skincare* di *E-Commerce*

Menyatakan dengan sebenarnya bahwa Skripsi yang saya tulis ini benar-benar merupakan hasil karya saya sendiri, bukan merupakan pengambil alihan data, tulisan, atau pikiran orang lain yang saya akui sebagai hasil tulisan atau pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar pustaka.

Apabila dikemudian hari terbukti atau dapat dibuktikan skripsi ini merupakan hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 15 Desember 2025
Yang membuat pernyataan,



Jinan Naurah
NIM. 210605110136

MOTTO

“Jangan biarkan rasa takut menghentikan langkahmu. Beranilah melangkah, karena keberanian lahir dari perjuangan, bukan dari kenyamanan.”

“Teruslah melangkah, meski ragu dan takut datang silih berganti. Sebab diam tak akan pernah membawamu ke mana pun.”

“Keberhasilan bukanlah milik orang pintar. Keberhasilan adalah kepunyaan mereka yang senantiasa berusaha.”

HALAMAN PERSEMBAHAN

Segala puji dan syukur penulis panjatkan ke hadirat Allah SWT atas limpahan rahmat, taufik, serta hidayah-Nya. Berkat pertolongan dan kemudahan dari-Nya, penulis dapat menyelesaikan skripsi ini dengan baik. Dengan penuh ketulusan dan rasa syukur, karya ini penulis persembahkan kepada:

Mama tercinta, Qowiyatul Aminah, serta Papa tercinta, Mohammad Fakhruddin, atas segala kasih sayang yang tak pernah pudar, doa yang tak pernah putus, serta dukungan yang selalu menguatkan di setiap langkah. Terima kasih atas kesabaran, pengorbanan, dan ketulusan yang menjadi sumber kekuatan terbesar dalam hidup penulis.

Keluarga besar penulis,
yang tidak pernah berhenti memberikan doa, perhatian, dan semangat. Setiap dukungan yang kalian berikan menjadi penguat tersendiri di tengah perjalanan panjang menuju terselesaikannya karya ini.

Dan untuk diri penulis sendiri,
Terima kasih telah kuat menghadapi setiap ujian, tetap berusaha ketika hampir menyerah, dan terus melangkah hingga akhir. Semoga perjuangan ini menjadi awal dari perjalanan panjang yang penuh makna dan keberhasilan.

KATA PENGANTAR

Assalamu'alaikum Warahmatullahi Wabarakatuh

Puji syukur kehadiran Allah SWT atas segala limpahan rahmat, taufik, dan hidayah-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul “Penerapan *Logistic Regression* Untuk Analisis Sentimen Pada Ulasan Produk *Skincare* di *E-Commerce*” dengan baik. Shalawat serta salam semoga senantiasa tercurah kepada junjungan Nabi Muhammad SAW yang telah membawa umat manusia dari zaman kegelapan menuju zaman yang terang benderang dengan ilmu pengetahuan.

Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer (S.Kom) pada Program Studi Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang. Dalam proses penyusunan skripsi ini, penulis banyak menerima bimbingan, bantuan, dan doa dari berbagai pihak. Oleh karena itu, dengan segala kerendahan hati penulis menyampaikan terima kasih kepada:

1. Prof. Dr. Hj. Ilfi Nur Diana, M.Si, selaku Rektor Universitas Islam Negeri Maulana Malik Ibrahim Malang.
2. Dr. Agus Mulyono, M.Kes, selaku Dekan Fakultas Sains dan Teknologi.
3. Supriyono, M.Kom, selaku Ketua Program Studi Teknik Informatika.
4. Prof. Dr. Muhammad Faisal, M.T, selaku Dosen Pembimbing I yang dengan penuh kesabaran, ketulusan, dan ketelitian telah memberikan bimbingan, arahan, serta motivasi kepada penulis selama penyusunan skripsi ini.

5. Dr. Ririen Kusumawati, S.Si, M.Kom, selaku Dosen Pembimbing II yang telah memberikan masukan dalam penyempurnaan penelitian ini.
6. Supriyono, M.Kom, selaku Dosen Penguji I, dan Hani Nurhayati, M.T, selaku Dosen Penguji II, atas saran, kritik, serta masukan yang membangun dalam penyempurnaan hasil penelitian ini.
7. Seluruh Dosen dan Staf Fakultas Sains dan Teknologi, khususnya Program Studi Teknik Informatika, yang telah memberikan ilmu, bimbingan, dan pelayanan terbaik selama masa perkuliahan.
8. Kedua orang tua tercinta, atas segala kasih sayang, doa, serta dukungan moral dan material yang tiada henti. Terima kasih atas setiap pengorbanan, kesabaran, dan kerja keras yang telah diberikan dengan penuh keikhlasan. Dukungan dan doa dari kedua orang tua menjadi kekuatan utama bagi penulis untuk terus berjuang dan tidak menyerah dalam menghadapi berbagai kesulitan selama proses perkuliahan hingga penyusunan skripsi ini. Semangat dan doa dari kedua orang tua senantiasa menjadi pengingat agar penulis tetap berusaha memberikan yang terbaik. Semoga Allah SWT membalas segala kasih sayang dan pengorbanan dengan kesehatan, kebahagiaan, dan keberkahan yang tiada henti. Terima kasih juga kepada kedua adik, yaitu Muhammad Nauvel Al-Abror dan Naureen Tsama Azka, atas doa, dukungan, serta semangat yang selalu diberikan. Kehadiran Naureen Tsama Azka turut menjadi penyemangat dan penghibur, terutama ketika merasa lelah, stres, maupun jenuh selama proses perkuliahan dan penyusunan skripsi ini.

9. Keluarga dan teman, yang telah memberikan perhatian, pendampingan, serta dukungan selama masa perkuliahan hingga penyusunan skripsi, khususnya Annisa Ayu Permadani yang dengan ketulusan dan kesetiaan telah menemani penulis dalam menyelesaikan berbagai tugas perkuliahan dan proses penyusunan skripsi, termasuk meluangkan waktu untuk mendampingi penulis belajar dan mengerjakan tugas dalam durasi yang panjang; serta Ratu Shabrina Nadhifa (sepupu penulis) yang senantiasa memberikan dukungan, perhatian, dan penguatan sehingga penulis dapat menjaga semangat dalam menyelesaikan studi dan skripsi ini.
10. Teman-teman yang telah kebersamai penulis dan memberikan dukungan dalam perjalanan perkuliahan hingga penyusunan skripsi ini, khususnya Salsabila Ramadanti Nuraini yang membantu menjawab pertanyaan penulis terkait perkuliahan; Agustina Mufidatuzzainiya yang berkenan berdiskusi, memberikan masukan, serta membantu menjawab pertanyaan penulis terkait skripsi; Farah Vairuza Natasya yang senantiasa memberikan dukungan dan ruang untuk berbagi; serta Lailatul Khoiriyah yang memberikan motivasi dalam pengerjaan skripsi dan membantu penulis memperoleh arahan awal selama proses penyusunan skripsi.
11. Teman-teman seperjuangan, khususnya Ferina Putri Soedjono, Vinka Rosalia Fatimah Devi, dan Nurul Aziza Kusuma Dewi, atas kebersamaan, dukungan, serta semangat yang diberikan selama masa perkuliahan.
12. Teman-teman Teknik Informatika angkatan 2021, khususnya teman-teman kelas E, terima kasih atas kebersamaan, semangat, dan dukungan selama

menempuh perjalanan perkuliahan. Semoga seluruh anggota ASTER senantiasa diberikan kesuksesan dan keberkahan dalam setiap langkah ke depan.

13. Seluruh pihak yang telah terlibat secara langsung maupun tidak langsung dalam proses penyusunan skripsi sejauh ini.

Penulis menyadari bahwa dalam penyusunan skripsi ini masih terdapat kekurangan, baik dari segi penulisan maupun isi. Oleh karena itu, penulis mengharapkan kritik dan saran yang membangun demi perbaikan di masa mendatang. Semoga skripsi ini dapat memberikan manfaat bagi pembaca dan menjadi kontribusi kecil dalam pengembangan ilmu pengetahuan.

Wassalamualaikum Warahmatullahi Wabarakatuh.

Malang, 15 Desember 2025

Penulis

DAFTAR ISI

HALAMAN PENGAJUAN	ii
HALAMAN PERSETUJUAN	iii
HALAMAN PENGESAHAN	iv
PERNYATAAN KEASLIAN TULISAN	v
MOTTO	vi
HALAMAN PERSEMBAHAN	vii
KATA PENGANTAR	viii
DAFTAR ISI	xii
DAFTAR GAMBAR	xiv
DAFTAR TABEL	xv
ABSTRAK	xvi
ABSTRACT	xvii
ملخص	xviii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	6
1.3 Batasan Masalah	6
1.4 Tujuan Penelitian	7
1.5 Manfaat Penelitian	7
BAB II STUDI PUSTAKA	8
2.1 Penelitian Terkait	8
2.2 Analisis Sentimen	11
2.3 Representasi Fitur dengan TF-IDF	12
2.4 Algoritma <i>Logistic Regression</i>	13
2.4.1 Model Multinomial <i>Logistic Regression</i>	15
2.5 Alasan Pemilihan <i>Logistic Regression</i>	16
2.6 Asumsi <i>Logistic Regression</i> dalam Klasifikasi Teks	17
2.7 Evaluasi Model	18
2.8 Ulasan Produk <i>Skincare</i> di <i>E-Commerce</i>	19
BAB III METODOLOGI PENELITIAN	21
3.1 Prosedur Penelitian	21
3.1.1 <i>Input</i> : Pengumpulan Data.....	23
3.1.2 <i>Process</i> : Representasi Fitur (TF-IDF)	26
3.1.3 <i>Process</i> : Pembagian Data.....	30
3.1.4 <i>Process</i> : Pemodelan <i>Logistic Regression</i>	32
3.1.5 <i>Process</i> : Perhitungan Evaluasi	38
3.1.6 <i>Output</i> : Hasil Evaluasi.....	41
3.1.7 <i>Pseudocode</i> Proses Klasifikasi Sentimen Menggunakan <i>Logistic Regression</i>	43
3.2 Skenario Pengujian	44
BAB IV HASIL DAN PEMBAHASAN	46
4.1 Deskripsi Umum Pengujian	46

4.2 Hasil Pemodelan <i>Logistic Regression</i>	46
4.3 Hasil Evaluasi	48
4.3.1 Skenario1	48
4.3.2 Skenario2.....	49
4.3.3 Skenario3.....	51
4.3.4 Skenario4.....	52
4.3.5 Perbandingan Antar Skenario	53
BAB V KESIMPULAN DAN SARAN.....	58
5.1 Kesimpulan	58
5.2 Saran	59
DAFTAR PUSTAKA	

DAFTAR GAMBAR

Gambar 3. 1 Alur Penelitian	22
Gambar 3. 2 Flowchart Proses Pemodelan <i>Logistic Regression</i>	37
Gambar 4. 1 <i>Confusion matrix</i> C=1_Split 80:20	49
Gambar 4. 2 <i>Confusion matrix</i> C=10_Split 80:20	50
Gambar 4. 3 <i>Confusion matrix</i> C=1_Split 70:30	52
Gambar 4. 4 <i>Confusion matrix</i> C=10_Split 70:30	53
Gambar 4. 5 Grafik Perbandingan Akurasi, <i>Precision</i> , <i>Recall</i> , dan F1-score Antar Skenario	54

DAFTAR TABEL

Tabel 2.1 Perbandingan Penelitian Terdahulu	9
Tabel 3. 1 Atribut Data pada Dataset <i>Skincare-product-review-indo-v2</i>	24
Tabel 3. 2 Distribusi Kelas Sentimen pada Dataset	25
Tabel 3. 3 Contoh Data Sebelum dan Sesudah Pembobotan TF-IDF	27
Tabel 3. 4 Nilai <i>Inverse Document Frequency</i> (IDF)	28
Tabel 3. 5 Contoh Gabungan Perhitungan TF-IDF	29
Tabel 3. 6 Contoh Matriks Hasil Pembobotan TF-IDF	29
Tabel 3. 7 Pembagian Data	32
Tabel 3. 8 Nilai TF-IDF (x) pada Dokumen D1	34
Tabel 3. 9 Nilai Koefisien (β) <i>Logistic Regression</i> (Ilustratif)	35
Tabel 3. 10 Contoh Perhitungan <i>Logistic Regression</i> – Dokumen D1	36
Tabel 3. 11 Struktur Umum <i>Confusion matrix</i>	39
Tabel 3. 12 Contoh <i>Confusion matrix</i> Hasil Klasifikasi <i>Logistic Regression</i> (Ilustratif)	39
Tabel 3. 13 Hasil Perhitungan Metrik Evaluasi (Contoh Manual)	41
Tabel 3. 14 Contoh Data Salah Klasifikasi (Ilustratif)	42
Tabel 3. 15 Skenario Pengujian	45
Tabel 4. 1 Hasil Klasifikasi Model <i>Logistic Regression</i>	47
Tabel 4. 2 Hasil Evaluasi C=1_Split 80:20	49
Tabel 4. 3 Hasil Evaluasi C=10_Split 80:20	50
Tabel 4. 4 Hasil Evaluasi setiap Kelas C=10_Split 80:20	51
Tabel 4. 5 Hasil Evaluasi C=1_Split 70:30	52
Tabel 4. 6 Hasil Evaluasi C=10_Split 70:30	53
Tabel 4. 7 Hasil Evaluasi Perbandingan 4 Skenario	54

ABSTRAK

Naurah, Jinan. 2025. **Penerapan *Logistic Regression* Untuk Analisis Sentimen Pada Ulasan Produk *Skincare* di *E-Commerce***. Skripsi. Program Studi Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (I) Prof. Dr. Muhammad Faisal, M.T (II) Dr. Ririen Kusumawati, S.Si, M.Kom.

Kata kunci: Analisis Sentimen, *TF-IDF*, *Logistic Regression*, Ulasan *Skincare*, *Machine Learning*.

Penelitian ini bertujuan untuk menerapkan metode TF-IDF dan *Logistic Regression* dalam mengklasifikasikan sentimen ulasan produk *Skincare* pada platform *E-Commerce*. Dataset yang digunakan merupakan dataset publik berbahasa Indonesia berisi tiga kategori sentimen, yaitu positif, netral, dan negatif. Tahapan penelitian meliputi pemuatan dataset, representasi teks menggunakan TF-IDF, pembagian data latih dan uji, serta pelatihan model *Logistic Regression* pada empat skenario pengujian dengan variasi parameter regularisasi ($C = 1$ dan $C = 10$) dan rasio pembagian data (80:20 dan 70:30). Hasil evaluasi menunjukkan bahwa seluruh skenario menghasilkan performa yang tinggi dengan nilai akurasi, *precision*, *recall*, dan *F1-score (macro)* berada di atas 90%. Model terbaik diperoleh pada skenario C10_80_20, yaitu *Logistic Regression* dengan parameter $C = 10$ dan rasio pembagian 80:20, yang mencapai akurasi tertinggi di antara seluruh skenario. Temuan ini menunjukkan bahwa kombinasi TF-IDF dan *Logistic Regression* mampu menghasilkan klasifikasi sentimen yang efektif dan dapat dijadikan dasar untuk pengembangan sistem analisis sentimen pada domain *E-Commerce*.

ABSTRACT

Naurah, Jinan. 2025. **Penerapan *Logistic Regression* Untuk Analisis Sentimen Pada Ulasan Produk *Skincare* di *E-Commerce***. Thesis. Informatics Engineering Study Program Faculty of Science and Technology Maulana Malik Ibrahim State Islamic University Malang. Advisors: (I) Prof. Dr. Muhammad Faisal, M.T (II) Dr. Ririen Kusumawati, S.Si, M.Kom.

This research aims to implement TF-IDF and *Logistic Regression* for classifying sentiment in *Skincare* product reviews on *E-Commerce* platforms. The study employs a public Indonesian-language dataset containing three sentiment categories: positive, neutral, and negative. The methodology includes dataset loading, text representation using TF-IDF, data splitting, and model training across four testing scenarios with variations in regularization parameter ($C = 1$ and $C = 10$) and data split ratios (80:20 and 70:30). The evaluation results indicate that all scenarios achieve high performance, with *accuracy*, *precision*, *recall*, and *macro F1-score* values exceeding 90%. The best-performing model is obtained in the C10_80_20 scenario, where *Logistic Regression* with $C = 10$ and an 80:20 train-test ratio achieves the highest *accuracy* among all configurations. These findings demonstrate that the combination of TF-IDF and *Logistic Regression* is effective for sentiment classification and can serve as a foundation for developing sentiment analysis systems in the *E-Commerce* domain.

Keywords: *Sentiment Analysis, TF-IDF, Logistic Regression, Skincare Reviews, Machine Learning*

ملخص

نورة، جنان. ٢٠٢٥. تطبيق الانحدار اللوجستي لتحليل المشاعر في مراجعات منتجة العناية بالبشرة على التجارة الإلكترونية. البحث الجامعي. قسم الهندسة المعلوماتية، كلية العلوم والتكنولوجيا بجامعة مولانا مالك إبراهيم الإسلامية الحكومية مالانج. المشرف الأول: أ. د. محمد فيصل، الماجستير؛ المشرف الثاني: د. ريرين كوسوماواتي، الماجستير.

الكلمات الرئيسية: تحليل مشاعر، $TF-IDF$ ، انحدار لوجستي، مراجعات عناية ببشرة، تعلم آلي

يهدف هذا البحث إلى تطبيق طريقة $TF-IDF$ والانحدار اللوجستي في تصنيف مشاعر مراجعات منتجة العناية بالبشرة على منصات التجارة الإلكترونية. مجموعة البيانات المستخدمة هي مجموعة بيانات عامة باللغة الإندونيسية وتحتوي على ثلاث فئات من المشاعر، وهي الإيجابية، والمحايدة، والسلبية. تشمل مراحل البحث تحميل مجموعة البيانات، وتمثيل النص باستخدام $TF-IDF$ ، وتقسيم البيانات إلى بيانات تدريب واختبار، وتدريب نموذج الانحدار اللوجستي على أربعة سيناريوهات اختبار مع اختلافات في معامل التنظيم ($ج=1$ و $ج=10$) ونسبة تقسيم البيانات ($80:20$ و $70:30$). أظهرت نتائج التقييم أن جميع السيناريوهات قدمت أداءً عاليًا مع قيمة الدقة، والضبط، والاستدعاء، وقيمة $F1$ (الإجمالي) أعلى من 90%. تم الحصول على أفضل نموذج في سيناريو $C10_80_20$ ، وهو الانحدار اللوجستي بمعامل $ج=10$ مع نسبة تقسيم $80:20$ ، والذي حقق أعلى دقة بين جميع السيناريوهات. أشارت هذه النتائج إلى أن الجمع بين $TF-IDF$ والانحدار اللوجستي قادر على إنتاج تصنيف فعال للمشاعر ويمكن أن يكون أساسًا لتطوير نظام تحليل المشاعر في مجال التجارة الإلكترونية.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi informasi telah membawa perubahan signifikan dalam berbagai aspek kehidupan, termasuk cara masyarakat memenuhi kebutuhan sehari-hari. Aktivitas belanja yang sebelumnya dilakukan secara konvensional kini beralih ke sistem daring melalui platform *E-Commerce* (Putra & Rahmawati, 2022). Kemudahan akses, keberagaman produk, serta efisiensi waktu dan biaya menjadikan *E-Commerce* sebagai salah satu sarana utama dalam aktivitas transaksi masyarakat Indonesia (Kementerian Perdagangan RI, 2023).

Seiring dengan meningkatnya penetrasi internet, tren belanja daring di Indonesia juga mengalami pertumbuhan yang pesat. Data *We Are Social* (2023) menunjukkan bahwa lebih dari 77% penduduk Indonesia telah menggunakan internet, dan sebagian besar di antaranya memanfaatkan platform *E-Commerce* untuk memenuhi kebutuhan konsumsi. Produk yang ditawarkan pun semakin beragam, tidak hanya mencakup kebutuhan rumah tangga dan elektronik, tetapi juga produk kecantikan, khususnya *Skincare*, yang mengalami peningkatan permintaan secara signifikan (Statista, 2023).

Produk *Skincare* menjadi salah satu kategori yang paling diminati seiring meningkatnya kesadaran masyarakat terhadap pentingnya perawatan kulit. Kondisi ini mendorong persaingan yang semakin ketat antar produsen *Skincare*, sehingga strategi pemasaran berbasis pengalaman dan opini konsumen menjadi faktor penting dalam memengaruhi keputusan pembelian (Romadhony dkk., 2024).

Dalam konteks ini, ulasan konsumen memiliki peran penting karena mencerminkan kepuasan, efektivitas produk, serta permasalahan yang dialami pengguna setelah menggunakan produk tertentu.

Keberadaan ulasan konsumen dalam platform *E-Commerce* memiliki pengaruh besar terhadap keputusan calon pembeli. Ulasan mampu memberikan gambaran mengenai kualitas produk, manfaat yang dirasakan, hingga keluhan pengguna (Fitria & Nugroho, 2020). Oleh karena itu, ulasan konsumen berfungsi sebagai *electronic word of mouth* (e-WOM) yang dipercaya sebagai referensi utama sebelum melakukan pembelian (Setiawan & Hidayat, 2021). Selain bermanfaat bagi konsumen, ulasan juga memiliki nilai strategis bagi produsen karena dapat digunakan sebagai dasar evaluasi produk, peningkatan kualitas layanan, serta penyusunan strategi pemasaran yang lebih tepat sasaran (Nugraha dkk., 2022).

Namun, jumlah ulasan yang sangat besar pada platform *E-Commerce* menimbulkan tantangan tersendiri. Ribuan hingga jutaan ulasan tidak memungkinkan untuk dianalisis secara manual karena membutuhkan waktu yang lama, biaya yang tinggi, serta berpotensi menimbulkan bias subjektivitas (Pratama & Wulandari, 2021). Oleh karena itu, diperlukan suatu pendekatan otomatis yang mampu mengolah dan menganalisis data teks ulasan secara cepat, akurat, dan konsisten.

Salah satu pendekatan yang banyak digunakan untuk mengatasi permasalahan tersebut adalah analisis sentimen. Analisis sentimen merupakan bagian dari text mining yang bertujuan untuk mengklasifikasikan opini atau sikap dalam teks ke dalam kategori positif, negatif, atau netral. Metode ini telah banyak

diterapkan dalam berbagai bidang, seperti evaluasi produk, pemantauan media sosial, serta analisis layanan publik (Zhang & Luo, 2020; Liu et al., 2022). Dalam konteks produk *Skincare*, analisis sentimen dapat membantu produsen dan konsumen memahami kecenderungan persepsi pasar secara sistematis dan berbasis data.

Agar teks ulasan dapat diproses oleh algoritma pembelajaran mesin, diperlukan representasi fitur yang mampu menggambarkan tingkat kepentingan kata dalam dokumen. Salah satu metode yang umum digunakan adalah *Term Frequency–Inverse Document Frequency* (TF-IDF), yaitu metode representasi teks yang mengubah dokumen menjadi vektor numerik dengan mempertimbangkan frekuensi kemunculan kata dalam dokumen serta tingkat kelangkaannya pada keseluruhan dataset (Agarwal & Mittal, 2021; Zhang et al., 2022). Representasi TF-IDF terbukti efektif dalam menangani data teks berdimensi tinggi dan banyak digunakan dalam penelitian analisis sentimen.

Algoritma klasifikasi yang digunakan dalam penelitian ini adalah *Logistic Regression*. *Logistic Regression* dipilih karena memiliki karakteristik yang sesuai untuk analisis sentimen berbasis teks hasil representasi TF-IDF, yaitu mampu menangani data berdimensi tinggi secara efisien serta menghasilkan performa klasifikasi yang stabil. Selain itu, *Logistic Regression* memiliki tingkat interpretabilitas yang baik, sehingga kontribusi kata-kata dalam menentukan kelas sentimen dapat dianalisis secara lebih jelas. Penelitian Zhang dan Luo (2020) menunjukkan bahwa *Logistic Regression* memiliki kinerja yang kompetitif pada klasifikasi teks berbasis TF-IDF. Selanjutnya, Romadhony dkk. (2024)

membuktikan bahwa *Logistic Regression* mampu bersaing dengan algoritma lain seperti *Support Vector Machine* dan *Naive Bayes* pada dataset ulasan *Skincare* berbahasa Indonesia. Penelitian terbaru oleh Özmen dan Gündüz (2025) juga menegaskan bahwa *Logistic Regression* merupakan model baseline yang efisien dan seimbang dalam analisis sentimen ulasan produk kecantikan berskala besar. Berdasarkan pertimbangan efisiensi komputasi, kemudahan interpretasi, serta dukungan hasil penelitian terkini, *Logistic Regression* dinilai tepat untuk digunakan dalam penelitian ini.

Meskipun penelitian analisis sentimen telah banyak dilakukan, sebagian besar penelitian di Indonesia masih berfokus pada domain umum atau produk elektronik. Penelitian yang secara khusus menganalisis ulasan produk *Skincare* masih relatif terbatas, padahal kategori ini memiliki jumlah ulasan yang besar dan berpengaruh terhadap perilaku konsumen. Hal ini menunjukkan adanya celah penelitian (*research gap*) yang perlu diisi melalui penelitian lebih lanjut (Wahyuni dkk., 2020).

Berdasarkan kondisi tersebut, penelitian ini memiliki kebaruan pada fokus domain dan penggunaan data. Kebaruan penelitian terletak pada analisis sentimen ulasan produk *Skincare* menggunakan dataset publik berbahasa Indonesia, yaitu *Skincare-product-review-indo-v2* yang tersedia di HuggingFace. Selain itu, penelitian ini menekankan penerapan kombinasi TF-IDF sebagai representasi teks dan *Logistic Regression* sebagai metode klasifikasi tanpa melibatkan algoritma pembanding lainnya, sehingga analisis dapat difokuskan secara mendalam pada evaluasi kinerja model.

Selain kontribusi teknis, penelitian ini juga relevan dengan nilai-nilai etika bisnis Islami. Islam menekankan pentingnya kejujuran, keterbukaan, dan keadilan dalam muamalah, termasuk dalam memberikan maupun memanfaatkan ulasan konsumen (Mufidah, 2021). Hal ini selaras dengan firman Allah SWT dalam Al-Qur'an:

يَا أَيُّهَا الَّذِينَ آمَنُوا إِن جَاءَكُمْ فَاسِقٌ بِنَبَأٍ فَتَبَيَّنُوا أَن تُصِيبُوا قَوْمًا بِجَهْلَةٍ فَتُصِبْحُوا عَلَىٰ مَا فَعَلْتُمْ نَادِمِينَ

“Wahai orang-orang yang beriman! Jika datang kepadamu orang fasik membawa suatu berita, maka periksalah dengan teliti agar kamu tidak menimpakan suatu musibah kepada suatu kaum karena kebodohan (kecerobohan) yang akhirnya kamu menyesali perbuatanmu itu.” (Qur'an 49:6).

Ayat tersebut menekankan prinsip *tabayyun*, yaitu memastikan kebenaran suatu informasi sebelum mempercayainya sebagai fakta. Dalam konteks ulasan digital produk *Skincare* di *E-Commerce*, prinsip ini menjadi sangat relevan mengingat tidak semua ulasan mencerminkan pengalaman konsumen yang objektif. Beberapa ulasan dapat bersifat bias, tidak akurat, atau bahkan direayasa. Oleh karena itu, analisis sentimen berbasis metode komputasional seperti *Logistic Regression* dapat dipandang sebagai upaya sistematis untuk mengolah informasi ulasan secara lebih objektif, terstruktur, dan dapat dipertanggungjawabkan.

Dengan demikian, penelitian ini diharapkan dapat memberikan kontribusi dalam pengembangan literatur akademik mengenai analisis sentimen berbasis TF-IDF dan *Logistic Regression*, serta memberikan manfaat praktis bagi produsen dan konsumen dalam memahami kecenderungan opini terhadap produk *Skincare* di *E-Commerce*. Berdasarkan uraian tersebut, penelitian ini difokuskan pada penerapan metode *Term Frequency –Inverse Document Frequency* (TF-IDF) dan algoritma

Logistic Regression dalam mengklasifikasikan sentimen ulasan produk *Skincare* di *E-Commerce*.

Dalam penelitian ini, metode klasifikasi yang digunakan adalah *Logistic Regression*. Pemilihan *Logistic Regression* didasarkan pada kesesuaiannya dengan karakteristik data teks hasil representasi TF-IDF yang berdimensi tinggi dan bersifat *sparse*. Selain itu, *Logistic Regression* memiliki interpretabilitas model yang baik serta efisiensi komputasi yang tinggi, sehingga banyak digunakan sebagai *strong baseline* dalam penelitian analisis sentimen. Oleh karena itu, *Logistic Regression* dipandang tepat untuk mengklasifikasikan sentimen ulasan produk *Skincare* di *E-Commerce* secara objektif dan sistematis.

1.2 Rumusan Masalah

Bagaimana penerapan metode *Term Frequency –Inverse Document Frequency* (TF-IDF) dengan algoritma *Logistic Regression* dalam mengklasifikasikan sentimen ulasan produk *Skincare* di *E-Commerce*?

1.3 Batasan Masalah

Ruang lingkup penelitian ini difokuskan pada penggunaan dataset publik “*Skincare-product-review-indo-v2*” dari HuggingFace yang berisi ulasan produk *Skincare* berbahasa Indonesia. Representasi teks dilakukan dengan metode *Term Frequency –Inverse Document Frequency (TF-IDF)* untuk mengubah data ulasan menjadi vektor numerik. Hasil representasi tersebut diproses menggunakan algoritma *Logistic Regression* sebagai metode klasifikasi sentimen. Penelitian ini

berfokus pada ulasan dalam domain *Skincare* di *E-Commerce* dan tidak melakukan perbandingan dengan algoritma lain di luar *Logistic Regression*.

1.4 Tujuan Penelitian

Tujuan penelitian ini untuk menerapkan model klasifikasi sentimen pada ulasan produk *Skincare* di *E-Commerce* dengan memanfaatkan metode *Term Frequency –Inverse Document Frequency* (TF-IDF) sebagai representasi teks dan algoritma *Logistic Regression* sebagai metode klasifikasi.

1.5 Manfaat Penelitian

Penelitian ini menghasilkan model klasifikasi sentimen pada ulasan produk *Skincare* di *E-Commerce* dengan menggunakan TF-IDF sebagai representasi teks dan *Logistic Regression* sebagai algoritma klasifikasi. Hasil penelitian ini diharapkan dapat memberikan manfaat, yaitu:

1. Memberikan gambaran mengenai kecenderungan opini konsumen terhadap produk *Skincare* di *E-Commerce* yang dapat dijadikan pertimbangan dalam evaluasi dan pengembangan produk.
2. Menjadi referensi bagi peneliti atau pihak terkait yang tertarik dalam pengembangan analisis sentimen berbasis teks menggunakan kombinasi metode representasi TF-IDF dan algoritma *Logistic Regression*.

BAB II

STUDI PUSTAKA

2.1 Penelitian Terkait

Penelitian mengenai analisis sentimen terhadap ulasan konsumen di *E-Commerce* telah banyak dilakukan dengan berbagai pendekatan. Wahyuni dkk. (2020) menganalisis sentimen ulasan produk di Tokopedia menggunakan *Naive Bayes* dan *Support Vector Machine* (SVM) dengan pendekatan leksikal, menghasilkan *accuracy* tertinggi 85% menggunakan SVM. Putra dkk. (2023) menggunakan SVM berbasis *Word Embedding* pada 2.500 ulasan produk elektronik di Shopee dengan *accuracy* 87%. Kusumaningrum & Nugroho (2021) menerapkan *Logistic Regression* dengan representasi TF-IDF pada 1.200 ulasan *E-Commerce* dengan *accuracy* 84%. Sementara itu, Amalia dkk. (2019) meneliti ulasan produk *fashion* menggunakan *Logistic Regression* berbasis TF-IDF dan memperoleh *accuracy* di atas 80%.

Sari dan Fadillah (2021) menganalisis sentimen ulasan produk kecantikan di Shopee menggunakan *Naive Bayes* pada 1.546 ulasan dengan *accuracy* 92,96%. Namun, sebagian besar penelitian tersebut fokus pada domain umum atau produk elektronik, sedangkan penerapan pada produk *Skincare* masih terbatas.

Penelitian terkini memberikan konteks yang lebih relevan. Özmen & Gündüz (2025) menganalisis 875 ribu ulasan kosmetik (*Skincare*, *body care*, dan lain sebagainya.) di Trendyol, membandingkan KNN, DT, RF, SVM, dan *Logistic Regression*. Hasilnya menunjukkan bahwa SVM unggul, namun *Logistic Regression* tetap kompetitif dengan keunggulan efisiensi komputasi dan kelemahan

di kelas netral. Nkiruka dkk. (2025) melakukan studi serupa pada ulasan pakaian dengan berbagai algoritma (LR, NB, SVM, RF, AdaBoost, GRU/LSTM), menemukan bahwa *Logistic Regression* seimbang dari sisi *accuracy* dan efisiensi. Romadhony dkk. (2024) memanfaatkan dataset *Skincare & makeup* Indonesia (>700.000 ulasan) dan menegaskan *Logistic Regression* sebagai *baseline* konvensional yang solid bersama SVM dan *Naive Bayes*.

Berbagai penelitian sebelumnya telah membahas penerapan metode pembelajaran mesin untuk analisis sentimen dengan domain dan algoritma yang berbeda-beda. Untuk memberikan gambaran umum mengenai pendekatan yang telah digunakan, Tabel 2.1 berikut merangkum beberapa penelitian terdahulu yang relevan dengan topik ini.

Tabel 2.1 Perbandingan Penelitian Terdahulu

Nama Peneliti	Judul Penelitian	Metode	Gap terhadap Penelitian Ini
Wahyuni dkk. (2020)	Analisis Sentimen Produk di Tokopedia Menggunakan <i>Naive Bayes</i> dan SVM	<i>Naive Bayes</i> , SVM (Leksikal)	Penelitian ini tidak fokus pada domain <i>Skincare</i> ; platform berbeda.
Putra dkk. (2023)	Sentiment Analysis <i>E-Commerce</i> Product Reviews Using Random Forest	SVM (<i>Word Embedding</i>)	Hanya menganalisis produk elektronik, bukan <i>Skincare</i>
Kusumaningrum & Nugroho (2021)	Sentiment Analysis Using <i>Logistic Regression</i> and <i>TF-IDF</i> on Consumer Review	<i>Logistic Regression</i> (<i>TF-IDF</i>)	Tidak mengkhususkan pada domain <i>Skincare</i> .
Amalia dkk. (2019)	Analisis Sentimen Ulasan Produk Fashion Menggunakan <i>Logistic Regression</i>	<i>Logistic Regression</i> (<i>TF-IDF</i>)	Fokus pada fashion, belum <i>Skincare</i> .
Sari & Fadillah (2021)	Analisis Sentimen Ulasan Produk Kecantikan di Shopee Menggunakan <i>Naive Bayes</i>	<i>Naive Bayes</i>	Menggunakan satu platform (Shopee), belum dataset gabungan.
Penelitian Ini (2025)	Penerapan <i>Logistic Regression</i> untuk Analisis Sentimen pada Ulasan Produk <i>Skincare</i> di <i>E-Commerce</i>	<i>Logistic Regression</i> (<i>TF-IDF</i>)	Berfokus pada domain <i>Skincare</i> Indonesia dengan dataset gabungan dari berbagai <i>E-Commerce</i> .

Berdasarkan penelusuran terhadap publikasi sebelumnya, belum ditemukan penelitian yang menggunakan dataset publik *Skincare-product-review-indo-v2* dari HuggingFace sebagai sumber data utama. Penelitian terdahulu umumnya menggunakan dataset yang dikumpulkan sendiri melalui *web scraping* dan terbatas pada satu platform *E-Commerce*. Dengan demikian, penelitian ini menjadi salah satu penelitian awal yang memanfaatkan dataset tersebut secara langsung, sehingga memiliki kebaruan pada sisi penggunaan dataset publik *Skincare* Indonesia.

Berdasarkan tabel di atas, penelitian ini diarahkan untuk mengisi kekosongan (*research gap*) melalui fokus pada domain *Skincare* di *E-Commerce* Indonesia secara umum. Berbeda dengan penelitian terdahulu yang umumnya terbatas pada Shopee atau platform tertentu, penelitian ini menggunakan dataset publik *Skincare-product-review-indo-v2* dari HuggingFace yang merupakan gabungan ulasan *Skincare* dari berbagai *E-Commerce*. Selain itu, penelitian ini menekankan penggunaan *Logistic Regression* dengan representasi TF-IDF sebagai metode utama, serta menyertakan evaluasi menyeluruh termasuk analisis kesalahan.

Berdasarkan penelitian-penelitian sebelumnya, dapat disimpulkan bahwa analisis sentimen pada ulasan *Skincare* masih relatif jarang dilakukan, terutama dengan memanfaatkan dataset gabungan dari berbagai *E-Commerce*. Oleh karena itu, penelitian ini berfokus pada penerapan representasi TF-IDF dengan algoritma *Logistic Regression* untuk menganalisis sentimen ulasan *Skincare* di *E-Commerce*, yang masih jarang dieksplorasi oleh penelitian terdahulu.

Berdasarkan hasil telaah literatur, tidak ditemukan penelitian dengan judul yang identik dengan penelitian ini. Namun terdapat beberapa penelitian yang memiliki kesamaan pada aspek metode maupun domain, meskipun belum secara spesifik membahas analisis sentimen pada ulasan produk *Skincare* Indonesia dengan menggunakan *Logistic Regression* sebagai model tunggal. Dengan demikian, penelitian ini menempati ruang kebaruan pada fokus domain *Skincare*, penggunaan dataset gabungan *E-Commerce*, serta penerapan *Logistic Regression* tanpa perbandingan algoritma lain.

2.2 Analisis Sentimen

Analisis sentimen adalah proses untuk mengidentifikasi, mengekstraksi, dan mengklasifikasikan opini subjektif dalam teks menjadi kategori positif, negatif, atau netral (Wibowo dkk., 2021). Dalam konteks *E-Commerce*, analisis sentimen memberikan wawasan mengenai tingkat kepuasan pelanggan, persepsi terhadap kualitas produk, dan potensi perbaikan layanan (Putra & Lestari, 2022).

Pendekatan analisis sentimen dapat dibagi menjadi dua: berbasis leksikon dan berbasis pembelajaran mesin (*machine learning*). Pendekatan leksikon menggunakan kamus kata yang memiliki bobot sentimen, namun lemah dalam menangkap konteks kompleks. Sementara itu, pendekatan berbasis machine learning lebih fleksibel dengan mengubah teks menjadi representasi numerik (misalnya menggunakan TF-IDF) sebelum diproses dengan algoritma klasifikasi seperti *Naive Bayes*, SVM, atau *Logistic Regression*.

Kategori sentimen yang umum digunakan adalah positif, negatif, dan netral. Kategori positif menunjukkan kepuasan, negatif menunjukkan kekecewaan,

sedangkan netral bersifat informatif. Tingkat analisis dapat dilakukan pada level dokumen, kalimat, atau aspek tertentu seperti harga, kemasan, dan tekstur (Sari & Hidayah, 2018).

Dalam penelitian ini, analisis sentimen digunakan untuk mengklasifikasikan opini konsumen terhadap produk *Skincare* ke dalam tiga kelas sentimen, yaitu positif, negatif, dan netral. Proses ini bertujuan untuk mengubah data ulasan teks yang bersifat tidak terstruktur menjadi informasi terstruktur yang dapat dianalisis secara kuantitatif menggunakan algoritma *machine learning*, khususnya *Logistic Regression*.

2.3 Representasi Fitur dengan TF-IDF

Dalam pengolahan data teks, diperlukan representasi numerik agar teks dapat diproses oleh algoritma machine learning. Salah satu pendekatan yang umum digunakan adalah *Term Frequency –Inverse Document Frequency* (TF-IDF), yang mengubah dokumen teks menjadi vektor fitur berdasarkan tingkat kepentingan kata.

Rumus TF-IDF adalah sebagai berikut:

$$TF_{t,d} = \frac{f_{t,d}}{\sum_k f_{k,d}} \quad (2.1)$$

$$IDF_t = \log \left(\frac{N}{df_t} \right) \quad (2.2)$$

$$TFIDF_{t,d} = TF_{t,d} \times IDF_t \quad (2.3)$$

Keterangan :

$f_{t,d}$	= jumlah kemunculan term t dalam dokumen d
$\sum_k f_{k,d}$	= total semua term pada dokumen d
N	= jumlah total dokumen

df_t	= jumlah dokumen yang mengandung kata t
$TF_{t,d}$	= frekuensi relatif term dalam dokumen
IDF_t	= tingkat kelangkaan term dalam korpus
$TFIDF_{t,d}$	= bobot kata hasil kombinasi TF dan IDF

Rumus (2.1) menghitung seberapa sering kata muncul pada suatu dokumen, (2.2) menghitung seberapa jarang kata muncul pada keseluruhan dokumen, dan (2.3) mengalikan keduanya untuk menentukan bobot penting setiap kata dalam representasi vektor.

2.4 Algoritma *Logistic Regression*

Logistic Regression merupakan algoritma klasifikasi yang banyak digunakan dalam analisis sentimen karena memiliki interpretabilitas model yang baik, efisiensi komputasi yang tinggi, serta kemampuan menangani data berdimensi tinggi. Karakteristik tersebut menjadikan *Logistic Regression* sesuai untuk diterapkan pada data teks yang direpresentasikan dalam bentuk vektor numerik, seperti hasil pembobotan TF-IDF.

Secara umum, *Logistic Regression* memodelkan probabilitas suatu data masuk ke dalam kelas tertentu dengan membentuk kombinasi linier dari fitur-fitur *Input*, kemudian memetakan hasilnya ke dalam rentang nilai probabilitas antara 0 dan 1 melalui fungsi logit atau sigmoid pada kasus klasifikasi biner, serta fungsi softmax pada kasus klasifikasi multi-kelas (Hosmer dkk., 2013). Kombinasi linier fitur dan fungsi sigmoid yang menjadi inti dari *Logistic Regression* dinyatakan sebagai berikut (Sutarman dkk., 2024):

$$z = \beta_0 + \sum_{i=1}^n \beta_i x_i \quad (2.4)$$

$$P(y = 1|x) = \sigma(z) = \frac{1}{1 + e^{-z}} \quad (2.5)$$

Keterangan :

z	= nilai kombinasi linier dari fitur
β_0	= bias (intercept)
β_i	= koefisien pembobot untuk fitur ke i
x_i	= nilai fitur TF – IDF ke i
n	= jumlah total fitur
$P(y = 1 x)$	= probabilitas bahwa teks termasuk kelas positif

Nilai probabilitas $P(y=1|x)$ yang dihasilkan oleh fungsi sigmoid digunakan sebagai dasar dalam pengambilan keputusan klasifikasi. Pada klasifikasi biner, suatu data dikategorikan ke dalam kelas target apabila nilai probabilitasnya melebihi ambang batas tertentu, yang umumnya ditetapkan sebesar 0,5. Sebaliknya, apabila nilai probabilitas berada di bawah ambang batas tersebut, maka data diklasifikasikan ke dalam kelas lainnya.

Dalam konteks analisis teks, *Logistic Regression* memiliki sejumlah keunggulan, antara lain interpretabilitas koefisien yang tinggi, efisiensi dalam proses pelatihan, serta kemampuan menangani data berdimensi tinggi ketika dikombinasikan dengan representasi TF-IDF (Zhang & Luo, 2020). Berbagai penelitian terdahulu menunjukkan bahwa *Logistic Regression* mampu memberikan performa yang kompetitif pada tugas analisis sentimen. Romadhony dkk. (2024) melaporkan bahwa *Logistic Regression* dapat bersaing dengan metode lain seperti *Support Vector Machine* dan *Naive Bayes* pada dataset ulasan *Skincare* berbahasa Indonesia. Selain itu, penelitian *Sentiment Classification of MyPertamina Reviews*

Using Naive Bayes and Logistic Regression (2022) menunjukkan bahwa *Logistic Regression* menghasilkan keseimbangan *precision* dan *recall* yang baik ketika menggunakan representasi TF-IDF. Penelitian *Logistic Regression Classification with TF-IDF and FastText for Sentiment Analysis of LinkedIn Reviews* (2024) juga menekankan efisiensi pelatihan sebagai salah satu keunggulan *Logistic Regression* dalam pengolahan data teks. Berdasarkan karakteristik tersebut, *Logistic Regression* dipandang sesuai untuk digunakan sebagai algoritma klasifikasi sentimen dalam penelitian ini.

2.4.1 Model Multinomial Logistic Regression

Penelitian ini menggunakan *Logistic Regression* untuk tugas klasifikasi multikelas, karena variabel target yang dianalisis memiliki lebih dari dua kategori sentimen, yaitu positif, negatif, dan netral. Oleh karena itu, model yang diterapkan adalah Multinomial *Logistic Regression*, yang merupakan pengembangan dari *Logistic Regression* biner untuk menangani lebih dari dua kelas.

Berbeda dengan klasifikasi biner yang menggunakan fungsi sigmoid, Multinomial *Logistic Regression* memanfaatkan fungsi *softmax* untuk menghitung probabilitas setiap kelas sentimen secara simultan. Probabilitas suatu dokumen teks x termasuk ke dalam kelas ke- k dirumuskan sebagai berikut:

$$P(y = k|x) = \frac{e^{\beta_k^T x}}{\sum_{j=1}^K e^{\beta_j^T x}} \quad (2.6)$$

Keterangan :

K = jumlah total kelas sentimen

β_k = vektor koefisien untuk kelas ke- k

x = vektor fitur hasil pembobotan TF-IDF

Setiap kelas sentimen memiliki vektor koefisien tersendiri yang merepresentasikan pengaruh masing-masing fitur terhadap peluang suatu dokumen termasuk ke dalam kelas tersebut. Hasil klasifikasi ditentukan berdasarkan kelas dengan nilai probabilitas tertinggi.

Pendekatan Multinomial *Logistic Regression* memungkinkan model untuk menangani klasifikasi sentimen positif, negatif, dan netral secara bersamaan dalam satu kerangka model. Dengan demikian, model ini sesuai untuk diterapkan pada penelitian ini yang berfokus pada analisis sentimen multikelas pada ulasan produk *Skincare* di *E-Commerce*.

2.5 Alasan Pemilihan *Logistic Regression*

Pemilihan *Logistic Regression* dalam penelitian ini didasarkan pada beberapa pertimbangan metodologis. Pertama, *Logistic Regression* mampu menangani data berdimensi tinggi yang dihasilkan oleh representasi TF-IDF secara stabil dan efisien. Kedua, model ini memiliki interpretabilitas yang baik karena koefisien regresi dapat digunakan untuk menganalisis kontribusi setiap fitur terhadap kelas sentimen.

Ketiga, *Logistic Regression* memiliki efisiensi komputasi yang tinggi dibandingkan algoritma klasifikasi yang lebih kompleks seperti *Support Vector Machine* berbasis kernel atau model *deep learning*, sehingga sesuai untuk penelitian dengan fokus evaluasi model secara mendalam. Keempat, berbagai penelitian menyebutkan *Logistic Regression* sebagai strong baseline dalam analisis sentimen dan klasifikasi teks.

Selain itu, penelitian ini tidak bertujuan membandingkan berbagai algoritma klasifikasi, melainkan memfokuskan analisis pada performa *Logistic Regression* dengan variasi parameter dan skenario pembagian data. Oleh karena itu, *Logistic Regression* dipandang sebagai metode yang tepat dan relevan dengan tujuan penelitian.

2.6 Asumsi *Logistic Regression* dalam Klasifikasi Teks

Logistic Regression merupakan algoritma klasifikasi yang memiliki sejumlah asumsi dasar yang perlu dipenuhi agar model dapat digunakan secara tepat. Asumsi-asumsi tersebut perlu dipahami dalam konteks data teks hasil representasi TF-IDF yang digunakan pada penelitian ini.

Pertama, *Logistic Regression* mensyaratkan bahwa variabel dependen bersifat kategorikal. Dalam penelitian ini, variabel target berupa kelas sentimen yang terdiri dari tiga kategori, yaitu positif, negatif, dan netral, sehingga telah memenuhi asumsi tersebut.

Kedua, *Logistic Regression* mengasumsikan adanya hubungan linier antara variabel independen dan log-odds dari probabilitas kelas. Hubungan linier ini tidak berlaku pada data asli, melainkan pada ruang logit. Oleh karena itu, meskipun data berasal dari teks, asumsi ini tetap terpenuhi ketika fitur direpresentasikan dalam bentuk numerik menggunakan TF-IDF.

Ketiga, model ini mengasumsikan bahwa setiap observasi bersifat independen. Setiap ulasan produk *Skincare* pada dataset dianggap sebagai satu dokumen yang berdiri sendiri dan tidak saling memengaruhi, sehingga asumsi independensi observasi terpenuhi.

Keempat, *Logistic Regression* tidak mensyaratkan distribusi normal pada data maupun residual. Hal ini menjadikan *Logistic Regression* sesuai untuk data teks yang umumnya tidak mengikuti distribusi tertentu dan memiliki karakteristik berdimensi tinggi.

Kelima, potensi multikolinearitas pada fitur TF-IDF dapat ditangani melalui mekanisme regularisasi. Pada penelitian ini, parameter regularisasi C digunakan untuk mengontrol kompleksitas model dan mencegah overfitting akibat korelasi antar fitur. Dengan demikian, asumsi *Logistic Regression* telah selaras dengan karakteristik data dan tujuan penelitian.

2.7 Evaluasi Model

Evaluasi model bertujuan mengukur kinerja klasifikasi sentimen secara menyeluruh. Metrik yang digunakan meliputi *accuracy*, *precision*, *recall*, dan *F1-score*. Metrik *accuracy* mengukur proporsi klasifikasi yang benar terhadap seluruh data uji dan dihitung menggunakan rumus:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2.7)$$

Keterangan :

<i>TP (True Positive)</i>	= jumlah data positif yang diklasifikasi positif
<i>TN (True Negative)</i>	= jumlah data negatif yang diklasifikasi negatif
<i>FP (False Positive)</i>	= jumlah data negatif yang diklasifikasi positif
<i>FN (False Negative)</i>	= jumlah data positif yang diklasifikasi negatif

Precision mengukur ketepatan klasifikasi positif:

$$Precision = \frac{TP}{TP + FP} \quad (2.8)$$

Recall mengukur seberapa banyak data positif yang berhasil dikenali:

$$Recall = \frac{TP}{TP + FN} \quad (2.9)$$

F1-score adalah rata-rata keseimbangan antara *precision* dan *recall*:

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (2.10)$$

Untuk dataset dengan ketidakseimbangan kelas, digunakan *macro F1-score* (rata-rata tanpa bobot) dan *weighted F1-score* (rata-rata berbobot sesuai jumlah data tiap kelas). Özmen & Gündüz (2025) dan Romadhony, dll (2024) merekomendasikan pelaporan keduanya untuk analisis sentimen dengan kelas netral.

2.8 Ulasan Produk *Skincare* di *E-Commerce*

E-Commerce seperti Shopee, Tokopedia, Lazada, dan Blibli merupakan platform yang menyediakan beragam produk, termasuk kategori *Skincare* yang terus mengalami peningkatan permintaan (DataReportal, 2023). Pada kategori *Skincare*, ulasan konsumen menjadi salah satu aspek penting yang memengaruhi keputusan pembelian. Ulasan tersebut biasanya singkat, menggunakan bahasa campuran (*code-switching*), serta sering mengandung emotikon.

Tantangan dalam menganalisis ulasan *Skincare* di *E-Commerce* meliputi: (1) ketidakseimbangan kelas sentimen, (2) dominasi kata-kata umum yang kurang informatif, dan (3) kecenderungan ulasan netral yang ambigu. Penelitian Najmi dkk. (2025) menunjukkan bahwa teknik penanganan ketidakseimbangan data seperti SMOTE dapat meningkatkan kinerja model klasifikasi.

Penelitian ini menggunakan dataset publik *Skincare-product-review-indo-v2* dari HuggingFace, yang berisi lebih dari 5.000 ulasan konsumen berbahasa Indonesia. Dataset ini berasal dari berbagai *E-Commerce*, sehingga tidak terbatas pada satu platform tertentu. Dengan karakteristik tersebut, dataset ini merepresentasikan keragaman ulasan *Skincare* di *E-Commerce* Indonesia, sekaligus memperkuat urgensi analisis sentimen untuk membantu konsumen dan produsen memahami kecenderungan opini pasar.

BAB III

METODOLOGI PENELITIAN

Bab ini menjelaskan tahapan metodologis yang digunakan dalam penelitian untuk membangun dan mengevaluasi model klasifikasi sentimen pada ulasan produk *Skincare* di *E-Commerce*. Penelitian ini dilakukan dengan pendekatan eksperimen komputasional yang mencakup proses pengumpulan data, pembentukan fitur menggunakan TF-IDF, penerapan algoritma *Logistic Regression*, serta evaluasi kinerja model menggunakan metrik klasifikasi. Setiap tahap dijelaskan secara sistematis sesuai kerangka *Input–Process–Output* (IPO), sehingga alur penelitian dapat dipahami secara menyeluruh dan terukur terhadap tujuan yang telah dirumuskan pada Bab I. Rumus-rumus dasar yang digunakan telah dijabarkan pada Bab II (Rumus 2.1 – 2.10); pada Bab III hanya disajikan implementasi numerik dan contoh perhitungan yang mengacu pada rumus-rumus tersebut.

3.1 Prosedur Penelitian

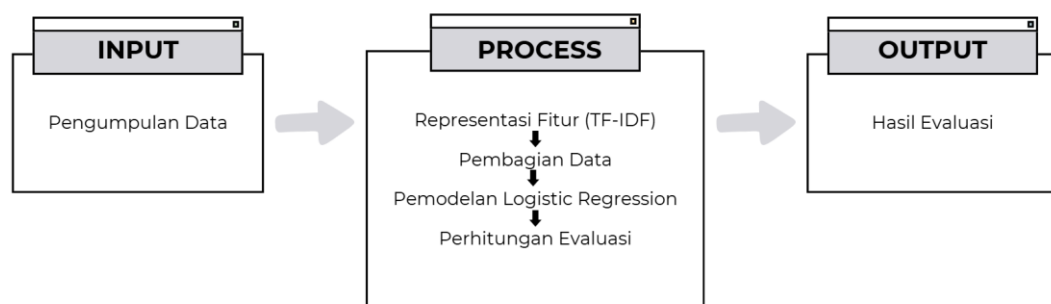
Prosedur penelitian ini disusun berdasarkan kerangka alur penelitian yang menggambarkan tahapan utama secara sistematis, mulai dari pengumpulan data hingga tahap evaluasi hasil klasifikasi.

Kerangka ini digunakan untuk memastikan setiap langkah penelitian berjalan secara terarah dan terukur sehingga seluruh tahapan memiliki keterkaitan yang jelas terhadap tujuan penelitian, yaitu membangun dan menguji model

klasifikasi sentimen menggunakan representasi fitur TF-IDF dan algoritma *Logistic Regression*.

Kerangka alur penelitian terdiri atas tiga komponen utama yang saling berkesinambungan, yaitu *Input*, *Process*, dan *Output*. Masing-masing komponen memiliki peran tersendiri namun saling berkaitan membentuk keseluruhan prosedur penelitian.

Secara umum, kerangka alur penelitian dalam studi ini dapat dilihat pada Gambar 3.1 berikut.



Gambar 3. 1 Alur Penelitian

Gambar tersebut menjelaskan bahwa penelitian ini diawali dengan tahap *Input*, yaitu pengumpulan data ulasan produk *Skincare* berbahasa dari dataset yang telah tersedia. Kolom *stem_reviews* digunakan sebagai *Input* teks dan kolom *sentimen* digunakan sebagai label klasifikasi. Data tersebut kemudian melalui tahap *Process*, data tersebut diubah menjadi representasi numerik menggunakan metode TF-IDF, kemudian dibagi menjadi data latih dan data uji. Selanjutnya, dilakukan pemodelan klasifikasi menggunakan algoritma *Logistic Regression* dan perhitungan evaluasi berdasarkan hasil klasifikasi model.

Tahap terakhir adalah *Output*, diperoleh hasil evaluasi kinerja model yang ditunjukkan melalui metrik *accuracy*, *precision*, *recall*, dan *F1-score* sebagai indikator keberhasilan model dalam mengklasifikasikan sentimen ulasan..

Kerangka penelitian ini memberikan gambaran menyeluruh mengenai arah dan tahapan penelitian secara umum tanpa menjelaskan detail teknis pada setiap tahap. Penjelasan yang lebih rinci mengenai masing-masing komponen akan diuraikan pada subbab berikutnya, dimulai dari tahap *Input*, yang mencakup deskripsi sumber dataset, atribut data, serta distribusi label sentimen yang digunakan dalam penelitian.

3.1.1 *Input* : Pengumpulan Data

Tahap *Input* merupakan langkah awal dalam penelitian ini yang berfungsi untuk menyiapkan data yang akan digunakan dalam proses pengolahan dan pemodelan. Data penelitian ini bersumber dari dataset publik berjudul “*Skincare-product-review-indo-v2*” yang tersedia pada platform HuggingFace Datasets. Dataset ini disusun oleh Shobrun Jibril dan berisi kumpulan ulasan produk *Skincare* yang ditulis oleh pengguna *E-Commerce* dalam bahasa Indonesia.

Dataset ini dipilih karena memiliki karakteristik yang sesuai dengan fokus penelitian, yaitu ulasan pendek dengan gaya bahasa sehari-hari yang merefleksikan opini pengguna terhadap produk *Skincare*. Karakteristik tersebut sangat relevan dengan tujuan penelitian ini, yaitu mengklasifikasikan sentimen pengguna terhadap produk *Skincare* berdasarkan isi teks ulasan.

Setiap baris data dalam dataset berisi sejumlah atribut yang digunakan untuk mendukung proses analisis. Secara umum, atribut-atribut penting yang digunakan dalam penelitian ini dapat dilihat pada Tabel 3.1 berikut.

Tabel 3. 1 Atribut Data pada Dataset *Skincare-product-review-indo-v2*

Nama Atribut	Deskripsi
<i>Username</i>	Nama pengguna yang memberikan ulasan.
<i>Rate</i>	Rating bintang yang diberikan pengguna terhadap produk.
<i>Date</i>	Tanggal ulasan diposting.
<i>Label</i>	Label kategori ulasan (positif, negatif, netral).
<i>Sentimen</i>	Sentimen ulasan sesuai anotasi dataset.
<i>Reviews</i>	Isi teks ulasan mentah dari pengguna.
<i>Produk</i>	Nama produk <i>Skincare</i> yang diulas.
<i>text_length</i>	Panjang teks ulasan dalam jumlah karakter.
<i>anomaly_score</i>	Skor anomali untuk mendeteksi ulasan yang tidak wajar
<i>lowercase reviews</i>	Ulasan dalam bentuk huruf kecil.
<i>no_emoji reviews</i>	Ulasan setelah emoji dihapus.
<i>no_hashtag reviews</i>	Ulasan setelah hashtag dihapus.
<i>no_number reviews</i>	Ulasan setelah angka dihapus.
<i>no_atribut reviews</i>	Ulasan setelah atribut tertentu dihapus.
<i>no_repetition reviews</i>	Ulasan setelah penghapusan huruf/karakter berulang
<i>no_whitespace reviews</i>	Ulasan setelah normalisasi spasi.
<i>no_banned reviews</i>	Ulasan setelah kata terlarang dihapus.
<i>no_slang reviews</i>	Ulasan setelah kata slang diganti dengan bentuk baku.
<i>clean reviews</i>	Ulasan hasil pembersihan menyeluruh.
<i>tokens reviews</i>	Ulasan setelah ditokenisasi menjadi kata.
<i>stopword reviews</i>	Ulasan setelah stopword dihapus.
<i>stem_reviews</i>	Ulasan setelah dilakukan <i>Stemming</i> .

Dari seluruh atribut tersebut, penelitian ini hanya memanfaatkan dua kolom utama, yaitu:

1. *stem_reviews* sebagai teks *Input* (fitur utama), dan
2. sentimen sebagai target *Output* (label kelas untuk pelatihan model).

Kolom-kolom lain seperti *username*, *rate*, dan produk digunakan hanya untuk keperluan verifikasi tambahan dan tidak dilibatkan dalam proses pemodelan klasifikasi sentimen.

Dataset ini berisi 6.000 ulasan produk *Skincare* berbahasa Indonesia yang telah diberi anotasi sentimen secara manual oleh penyusun dataset. Setiap kategori

sentimen—positif, negatif, dan netral—memiliki jumlah yang seimbang, yaitu masing-masing 2.000 ulasan. Namun, sebelum digunakan pada tahap pemodelan, dataset melalui proses pembersihan untuk menghapus nilai kosong pada kolom sentimen dan *stem_reviews*, menghapus ulasan kosong, serta memastikan konsistensi struktur data. Setelah proses pembersihan tersebut, jumlah data valid yang digunakan pada pemodelan menjadi 5.982 ulasan, terdiri dari 1.987 ulasan positif, 2.000 ulasan negatif, dan 1.995 ulasan netral. Ringkasan distribusi data mentah dan data setelah pembersihan dapat dilihat pada Tabel 3.2.

Tabel 3. 2 Distribusi Kelas Sentimen pada Dataset

Kelas Sentimen	Data Mentah	Data Setelah <i>Cleaning</i>	<i>Presentase</i>
Positif	2.000	1.987	33.3 %
Negatif	2.000	2.000	33.3 %
Netral	2.000	1.995	33.3 %
Total	6.000	5.982	100 %

Perbedaan jumlah antara data mentah dan data setelah pembersihan disebabkan oleh adanya nilai kosong, ulasan yang tidak memuat teks, serta beberapa entri yang tidak memenuhi kriteria kelayakan data. Proses pembersihan ini penting dilakukan agar model tidak belajar dari data yang tidak valid dan agar hasil klasifikasi yang diperoleh lebih akurat serta mencerminkan kondisi sebenarnya.

Persentase pada Tabel 3.2 diperoleh dari hasil analisis awal terhadap dataset publik *shobrunjb/Skincare-product-review-indo-v2* yang digunakan dalam penelitian ini. Dataset tersebut memiliki distribusi label yang seimbang, sehingga setiap kelas sentimen memiliki proporsi yang sama. Kondisi ini memberikan keuntungan pada tahap pelatihan model karena dapat mengurangi risiko bias

klasifikasi terhadap salah satu kelas sentimen, sehingga model dapat belajar secara proporsional dari masing-masing kategori.

Dalam konteks penelitian ini, bias mengacu pada kecenderungan model untuk lebih sering mengklasifikasi kelas tertentu akibat ketidakseimbangan data. Oleh karena itu, keseimbangan jumlah data pada setiap kelas dianggap penting agar hasil klasifikasi yang dihasilkan bersifat objektif dan representatif terhadap keseluruhan data ulasan.

Sebelum digunakan pada tahap pemodelan, dataset diperiksa untuk memastikan kualitas dan konsistensi data, meliputi:

1. Validasi struktur kolom, memastikan seluruh atribut terdefinisi dengan benar sesuai metadata dataset.
2. Pengecekan nilai kosong (*missing value*), terutama pada kolom sentimen dan *stem_reviews*.
3. Identifikasi data duplikat untuk memastikan setiap entri mewakili ulasan unik.
4. Seleksi label valid, yaitu hanya data dengan kategori positif, negatif, dan netral yang digunakan untuk proses pemodelan.

Tahap pengumpulan dan pemeriksaan data ini menjadi fondasi penting untuk tahap selanjutnya, yaitu proses ekstraksi fitur menggunakan TF-IDF, pembagian data latih dan uji, serta pemodelan klasifikasi menggunakan *Logistic Regression* untuk menganalisis sentimen ulasan produk *Skincare*.

3.1.2 Process: Representasi Fitur (TF-IDF)

Dataset telah disediakan dalam bentuk teks yang telah diproses oleh penyusun dataset, penelitian ini langsung menggunakan kolom *stem_reviews*

sebagai *Input*. Langkah berikutnya adalah mengubah data teks tersebut menjadi bentuk numerik agar dapat diproses oleh algoritma *Logistic Regression*. Representasi numerik yang digunakan dalam penelitian ini adalah *Term Frequency –Inverse Document Frequency* (TF-IDF).

Metode TF-IDF memberikan bobot numerik pada setiap kata (*term*) berdasarkan dua faktor utama, yaitu seberapa sering kata muncul dalam suatu ulasan (TF), dan seberapa jarang kata tersebut muncul di seluruh kumpulan data ulasan (IDF). Dengan kombinasi kedua komponen tersebut, sistem dapat menekankan kata-kata yang penting dalam konteks tertentu tetapi jarang muncul pada keseluruhan dataset. Dengan demikian, representasi TF-IDF mampu menghasilkan fitur numerik yang merefleksikan tingkat kepentingan kata dalam konteks sentimen.

Tabel 3. 3 Contoh Data Sebelum dan Sesudah Pembobotan TF-IDF

No	Teks Ulasan Setelah <i>Stemming</i>	Label Sentimen
1	banyak juang jerawat cocok face wash acne kulit produk kurang efektif timbul rasa perih pakai produk murah meriah	Negatif
2	sedih perih mata nangis bersih wajah bikin kulit kering jerawat kurang ngaruh kontrol harga murah beli supermarket rekomend mata sensitive	Negatif
3	facewash harga bagus jelek emang bikin tarik kulit separah suka coba jerawat strip kulit	Netral
4	bonus mantul scarlet baik	Positif
5	bener banyak cabang target middle low variasi harga buat tinggi demografi indonesia dominasi remaja menengah akhir mungkin produk beli kalangan rentang miskin	Netral
6	sayang tahan lama makan minum dikit langsung hilang padahal kemas oke produk kurang worth huhu	Negatif

1. *Term Frequency* (TF)

Tahapan pertama pada representasi fitur adalah menghitung *Term Frequency* (TF), yang mengukur seberapa sering suatu kata muncul dalam satu dokumen. Rumus dasar TF telah dijelaskan pada Rumus (2.1) di Bab II.

Sebagai ilustrasi implementasi, pada ulasan ke-1 (“banyak juang jerawat cocok *face wash acne* kulit produk kurang efektif timbul rasa perih pakai produk murah meriah”) jumlah total kata adalah 18.

Kemunculan kata “produk” sebanyak 2 kali, sehingga:

$$TF_{produk,d1} = \frac{2}{18} = 0,111 \quad (3.1)$$

Sedangkan kata lain yang muncul satu kali memiliki:

$$TF_{term,d1} = \frac{1}{18} = 0,055 \quad (3.2)$$

2. *Inverse Document Frequency (IDF)*

Langkah selanjutnya menghitung *Inverse Document Frequency* (IDF) untuk mengetahui seberapa penting kata tersebut di seluruh dokumen. Kata yang jarang muncul di banyak dokumen dianggap lebih informatif. Rumus IDF telah dijelaskan pada Rumus (2.2) di Bab II, yang menilai tingkat kelangkaan suatu kata pada keseluruhan kumpulan ulasan.

Tabel 3. 4 Nilai *Inverse Document Frequency* (IDF)

Kata	Dokumen yang Mengandung (df _i)	IDF = $\log(6/df_i)$
Produk	3	$\log(6/3) = 0,301$
Jerawat	3	$\log(6/3) = 0,301$
Kulit	3	$\log(6/3) = 0,301$
Murah	2	$\log(6/2) = 0,477$
Perih	2	$\log(6/2) = 0,477$
Efektif	1	$\log(6/1) = 0,778$
Worth	1	$\log(6/1) = 0,778$
Scarlet	1	$\log(6/1) = 0,778$
Middle	1	$\log(6/1) = 0,778$

Dari tabel di atas terlihat bahwa kata “efektif”, “worth”, “scarlet”, “middle” memiliki IDF paling tinggi karena hanya muncul di satu dokumen. Hal ini

menandakan bahwa kata-kata tersebut sangat spesifik dan informatif dalam konteks ulasan yang bersangkutan.

3. TF-IDF

Tahapan terakhir adalah menghitung bobot akhir setiap kata menggunakan metode TF-IDF. Rumus dasar TF-IDF yang merupakan hasil perkalian antara nilai TF dan IDF telah dijelaskan pada Rumus (2.3) di Bab II.

Tabel 3. 5 Contoh Gabungan Perhitungan TF-IDF

Kata	Tempat Muncul	TF	IDF	TF-IDF = TFxIDF
Produk	D1	0,111	0,301	0,033
Jerawat	D1	0,055	0,301	0,017
Kulit	D1	0,055	0,301	0,017
Murah	D1	0,055	0,477	0,026
Perih	D1	0,055	0,477	0,026
Efektif	D1	0,055	0,778	0,043
Worth	D6	0,059	0,778	0,046
Scarlet	D4	0,250	0,778	0,195
Middle	D5	0,041	0,778	0,032

Contoh implementasi rumus (untuk kata efektif pada dokumen ke-1) :

$$TFIDF_{efektif,d1} = 0,055 \times 0,778 = 0,043 \quad (3.3)$$

Setelah melalui proses perhitungan bobot menggunakan metode TF-IDF, seluruh teks ulasan diubah menjadi bentuk numerik yang merepresentasikan keterkaitan antara kata dan dokumen. Setiap baris dalam matriks menunjukkan satu dokumen (ulasan), sedangkan setiap kolom menunjukkan kata yang telah diberikan bobot berdasarkan tingkat kemunculannya dan kelangkaannya. Contoh hasil pembobotan ditunjukkan pada Tabel berikut.

Tabel 3. 6 Contoh Matriks Hasil Pembobotan TF-IDF

Kata	D1	D2	D3	D4	D5	D6
Produk	0,033	0,000	0,000	0,000	0,012	0,000
Jerawat	0,017	0,012	0,008	0,000	0,000	0,000
Kulit	0,017	0,014	0,007	0,000	0,000	0,000
Murah	0,026	0,020	0,000	0,000	0,000	0,000

Kata	D1	D2	D3	D4	D5	D6
Perih	0,026	0,018	0,000	0,000	0,000	0,000
Efektif	0,043	0,000	0,000	0,000	0,000	0,000
Worth	0,000	0,000	0,000	0,000	0,000	0,046
Scarlet	0,000	0,000	0,000	0,195	0,000	0,000
Middle	0,000	0,000	0,000	0,000	0,032	0,000

Hubungan Antar Dokumen dan Perhitungan TF-IDF

Setiap dokumen (D1–D6) pada tabel di atas mewakili enam ulasan yang diambil dari dataset penelitian. Perhitungan TF-IDF dilakukan secara individual untuk setiap dokumen, karena setiap ulasan dianggap sebagai satu unit analisis yang berdiri sendiri. Nilai TF dihitung berdasarkan frekuensi kemunculan kata dalam satu ulasan, sedangkan IDF mencerminkan kelangkaan kata di seluruh dataset.

Contoh perhitungan manual hanya ditampilkan untuk sebagian dokumen (D1, D4, D5, dan D6) yang memiliki kata-kata unik seperti “efektif”, “scarlet”, “middle”, dan “worth” agar representatif terhadap variasi konteks ulasan.

Pada implementasi sistem, seluruh dokumen (D1–D6) digunakan untuk membentuk matriks TF-IDF secara menyeluruh, yang kemudian menjadi *Input* bagi tahap pembagian data latih dan data uji sebelum dilakukan proses pemodelan *Logistic Regression*.

3.1.3 Process: Pembagian Data

Tahapan pembagian data bertujuan untuk memisahkan data hasil representasi fitur TF-IDF ke dalam dua bagian, yaitu data latih (*training data*) dan data uji (*testing data*). Data latih digunakan oleh model *Logistic Regression* untuk mempelajari pola hubungan antara fitur teks dan kelas sentimen, sedangkan data uji digunakan untuk mengevaluasi kemampuan model dalam melakukan klasifikasi

terhadap data baru yang belum pernah dilihat sebelumnya. Pemisahan kedua jenis data ini penting untuk memastikan bahwa evaluasi model dilakukan secara objektif dan tidak bias terhadap data pelatihan.

Dataset yang digunakan dalam penelitian ini merupakan dataset hasil pembersihan (*Cleaning*), yaitu sebanyak 5.982 ulasan valid. Jumlah ini diperoleh setelah proses penghapusan data yang memiliki nilai kosong (*missing value*) pada kolom *stem_reviews* dan sentimen, serta penghapusan ulasan yang tidak berisi teks. Kolom *stem_reviews* kemudian digunakan sebagai *Input* untuk proses ekstraksi fitur menggunakan TF-IDF, sedangkan kolom sentimen digunakan sebagai label pada proses pembelajaran model.

Pembagian data dilakukan menggunakan dua rasio, yaitu 80:20 dan 70:30. Variasi rasio ini digunakan untuk menilai konsistensi performa model *Logistic Regression* pada proporsi data latih yang berbeda. Proses pembagian data menerapkan teknik *stratified sampling* agar distribusi kelas sentimen (positif, netral, dan negatif) tetap proporsional antara data latih dan data uji. Secara umum, pembagian data dirumuskan sebagai berikut:

$$\begin{aligned} \text{Proporsi Data Latih} &= 0,8 \times N \\ \text{Proporsi Data Uji} &= 0,2 \times N \end{aligned} \tag{3.4}$$

Keterangan :

$$N = \text{jumlah total data ulasan}$$

Dengan jumlah total data sebanyak 5.982, maka perhitungannya adalah:

Skenario 80:20

$$\text{Proporsi Data Latih} = 0,8 \times 6000 = 4785 \tag{3.5}$$

$$\text{Proporsi Data Uji} = 0,2 \times 6000 = 1197$$

Skenario 70:30

$$\text{Proporsi Data Latih} = 0,8 \times 6000 = 4187 \quad (3.6)$$

$$\text{Proporsi Data Uji} = 0,2 \times 6000 = 1197$$

Ringkasan pembagian data untuk kedua skenario ditunjukkan pada Tabel

3.7.

Tabel 3. 7 Pembagian Data

Rasio Split	Data Latih	Data Uji	Total Data
80:20	4.785	1.197	5.982
70:30	4.187	1.795	5.982

Penggunaan dua variasi rasio pembagian data ini didasarkan pada praktik umum dalam penelitian analisis sentimen berbasis pembelajaran mesin, di mana proporsi data latih sebesar 70–80% dianggap optimal untuk mencapai keseimbangan antara kemampuan pembelajaran model dan daya generalisasi. Dengan pembagian data yang proporsional dan *stratified*, model diharapkan mampu mempelajari pola sentimen secara seimbang dan mengurangi potensi bias terhadap salah satu kelas sentimen. Tahapan selanjutnya setelah pembagian data adalah proses pemodelan menggunakan algoritma *Logistic Regression*, yang akan dijelaskan pada subbab berikutnya.

3.1.4 Process : Pemodelan *Logistic Regression*

Model klasifikasi yang digunakan pada penelitian ini adalah Multinomial *Logistic Regression*, karena variabel target terdiri dari tiga kelas sentimen, yaitu positif, negatif, dan netral. Data teks yang telah direpresentasikan dalam bentuk

vektor numerik menggunakan TF-IDF selanjutnya digunakan sebagai *Input* pada model *Logistic Regression*.

Dalam proses pemodelan, setiap dokumen ulasan direpresentasikan sebagai vektor fitur TF-IDF. Model *Logistic Regression* kemudian menghitung kombinasi linier antara vektor fitur dan koefisien regresi untuk setiap kelas sentimen. Nilai tersebut dipetakan menjadi probabilitas kelas menggunakan fungsi softmax, dan kelas dengan probabilitas tertinggi dipilih sebagai hasil klasifikasi sentimen.

Tahap pemodelan merupakan inti dari proses klasifikasi sentimen, di mana hasil pembobotan fitur TF-IDF digunakan sebagai *Input* ke dalam algoritma *Logistic Regression*. Algoritma ini dipilih karena sederhana, efisien, dan terbukti efektif dalam menangani data teks berukuran besar yang telah direpresentasikan dalam bentuk numerik.

Model *Logistic Regression* digunakan untuk mengklasifikasi kecenderungan sentimen suatu ulasan produk *Skincare* berbahasa Indonesia berdasarkan nilai fitur TF-IDF yang dihasilkan dari tahap sebelumnya.

Prinsip Dasar *Logistic Regression*

Proses perhitungan matematis *Logistic Regression* telah dijelaskan pada Bab II (Rumus 2.4 dan 2.5). Rumus (2.4) menjelaskan kombinasi linear antara fitur (hasil TF-IDF) dengan koefisien pembobotnya, sedangkan Rumus (2.5) menggambarkan konversi hasil kombinasi linear tersebut menjadi probabilitas melalui fungsi sigmoid.

Dengan demikian, pada subbab ini penjelasan berfokus pada implementasi praktis dan contoh perhitungan, bukan pada definisi matematis ulang. Langkah umum proses pemodelan adalah sebagai berikut:

1. Menggunakan hasil representasi fitur TF-IDF sebagai masukan (*Input matrix*).
2. Menginisialisasi parameter model (koefisien β dan *Intercept* β_0).
3. Menghitung kombinasi linear setiap fitur menggunakan Rumus (2.4).
4. Mengonversi hasilnya menjadi probabilitas menggunakan fungsi sigmoid sesuai Rumus (2.5).
5. Menetapkan ambang batas (threshold) 0,5 untuk menentukan kelas sentimen (Negatif atau Non-Negatif).

Nilai TF-IDF (x) yang Digunakan

Nilai TF-IDF yang digunakan diambil dari hasil pembobotan pada dokumen D1 yang merupakan salah satu contoh ulasan dalam dataset shobrunjb/*Skincare-product-review-indo-v2*. Nilai-nilai berikut menunjukkan bobot kata setelah tahap pembobotan TF-IDF dilakukan.

Tabel 3. 8 Nilai TF-IDF (x) pada Dokumen D1

Kata	TF-IDF (x)
Efektif	0,043
Perih	0,026
produk	0,033
Jerawat	0,017
Kulit	0,017
Murah	0,026

Nilai Koefisien (β)

Nilai koefisien di bawah ini digunakan secara ilustratif untuk menjelaskan langkah perhitungan *Logistic Regression* secara manual. Nilai-nilai sebenarnya diperoleh melalui proses pelatihan model yang akan dijelaskan pada Bab 4.

Tabel 3. 9 Nilai Koefisien (β) *Logistic Regression* (Ilustratif)

Fitur	β (Koefisien Ilustratif)	Alasan Pemilihan
efektif	8,0	Menunjukkan indikasi sentimen negatif kuat (contoh: “kurang efektif”)
perih	6,0	Kata negatif kuat (menunjukkan rasa tidak nyaman)
produk	0,2	Kata netral, umum pada semua sentimen
jerawat	0,7	Kata cukup berpengaruh pada konteks <i>Skincare</i>
kulit	0,5	Umum, relevan namun tidak dominan
murah	-0,2	Cenderung positif ringan
<i>Intercept</i> (β_0)	-0,2	Nilai bias awal model

Nilai koefisien (β) pada tabel di atas bersifat ilustratif untuk menunjukkan tahapan pengimplementasian rumus *Logistic Regression* secara manual. Nilai sebenarnya diperoleh melalui proses pelatihan model terhadap data latih yang telah ditentukan pada tahap sebelumnya (3.2.4), dan hasil pastinya akan ditampilkan serta dibahas pada Bab 4.

Pengimplementasian Rumus *Logistic Regression*

Perhitungan dilakukan dengan memanggil Rumus (2.4) untuk menghitung nilai kombinasi linear (z), dan Rumus (2.5) untuk menghitung probabilitas keluaran.

Sebagai contoh penerapan rumus *Logistic Regression*, digunakan data hasil pembobotan TF-IDF dari dokumen D1 (Tabel 3.9). Setiap nilai fitur x_i dikalikan dengan koefisien β_i , kemudian dijumlahkan bersama dengan nilai *Intercept* β_0 untuk menghasilkan nilai linear z .

Dengan menggunakan nilai ilustratif berikut:

$$\beta_0 = -0,2, \beta_{\text{efektif}} = 8,0, \beta_{\text{perih}} = 6,0, \beta_{\text{produk}} = 0,2, \beta_{\text{jerawat}} = 0,7, \beta_{\text{kulit}} = 0,5, \beta_{\text{murah}} = -0,2$$

dan nilai TF-IDF dari D1:

$$x_{\text{efektif}} = 0,043, x_{\text{perih}} = 0,026, x_{\text{produk}} = 0,033, x_{\text{jerawat}} = 0,017, x_{\text{kulit}} = 0,017, x_{\text{murah}} = 0,026$$

Maka perhitungan nilai linear z dapat dilakukan sebagai berikut:

$$z = (-0,2) + (8,0 \times 0,043) + (6,0 \times 0,026) + (0,2 \times 0,033) + (0,7 \times 0,017) + (0,5 \times 0,017) + (-0,2 \times 0,026)$$

$$z = -0,2 + 0,344 + 0,156 + 0,0066 + 0,0119 + 0,0085 - 0,0052 = 0,3218$$

Selanjutnya, menggunakan Rumus 2.5 yaitu nilai z dimasukkan ke dalam fungsi sigmoid untuk menghitung probabilitas:

$$P(y = 1|x) = \frac{1}{1 + e^{-z}} = \frac{1}{1 + e^{-0,3218}} = \frac{1}{1 + 0,7249} = 0,5798 \quad (3.7)$$

Karena nilai probabilitas $0,5798 \geq 0,5$, maka ulasan D1 diklasifikasikan sebagai sentimen negatif (kelas 1).

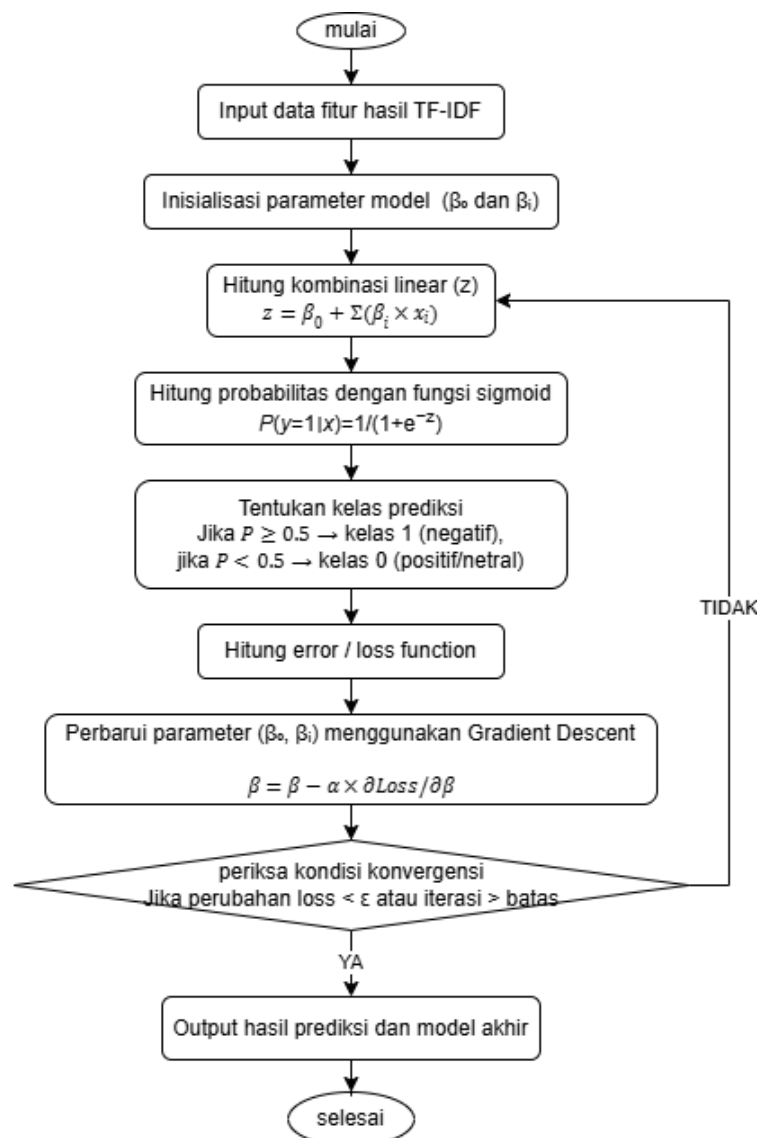
Tabel 3. 10 Contoh Perhitungan *Logistic Regression* – Dokumen D1

Fitur	TF-IDF (x)	Koefisien (β)	$\beta \times x$
Efektif	0,043	8,0	0,3440
Perih	0,026	6,0	0,1560
Produk	0,033	0,2	0,0066
Jerawat	0,017	0,7	0,0119
Kulit	0,017	0,5	0,0085
Murah	0,026	-0,2	-0,0052
Jumlah kontribusi fitur			0,5218
<i>Intercept</i> (β_0)		-0,2000	-0,2000
$z = \beta_0 + \sum \beta x$			0,3218
$\sigma(z)$ (probabilitas)			0,5798 (57,98%)
Klasifikasi			Negatif (kelas 1)

Tahapan pemodelan ini menjadi bagian penting dalam proses penelitian karena pada tahap inilah sistem “belajar” dari data untuk mengenali pola kata yang berkaitan dengan sentimen tertentu. Model *Logistic Regression* yang terbentuk akan digunakan pada tahap berikutnya, yaitu evaluasi kinerja model, untuk mengukur tingkat keberhasilan sistem dalam mengklasifikasikan sentimen pada ulasan produk *Skincare*.

Flowchart Proses *Logistic Regression*

Untuk memperjelas alur kerja algoritma yang digunakan, proses pemodelan *Logistic Regression* dalam penelitian ini divisualisasikan melalui flowchart pada Gambar 3.2 berikut. Flowchart ini menunjukkan tahapan perhitungan mulai dari *Input* fitur hasil TF-IDF hingga keluaran berupa klasifikasi kelas sentimen.



Gambar 3. 2 Flowchart Proses Pemodelan *Logistic Regression*

Berdasarkan alur pada Gambar 3.2, dapat dilihat bahwa proses *Logistic Regression* melibatkan tahapan iteratif dalam penyesuaian parameter hingga mencapai kondisi konvergensi. Model yang telah terbentuk selanjutnya akan dievaluasi untuk mengukur tingkat akurasi dan performa klasifikasi, sebagaimana dijelaskan pada subbab berikutnya.

3.1.5 *Process* : Perhitungan Evaluasi

Tahapan evaluasi kinerja model bertujuan untuk menilai sejauh mana algoritma *Logistic Regression* mampu mengklasifikasikan sentimen dengan baik terhadap data uji yang telah dipisahkan dari data latih. Evaluasi dilakukan untuk memastikan bahwa model tidak hanya mampu mengenali pola dari data pelatihan, tetapi juga dapat melakukan klasifikasi yang akurat pada data baru yang belum pernah dilihat sebelumnya.

Pada penelitian ini, proses evaluasi dilakukan menggunakan empat metrik utama, yaitu *Accuracy*, *Precision*, *Recall*, dan *F1-Score*. Keempat metrik tersebut dipilih karena mampu menggambarkan performa model secara menyeluruh, baik dari sisi ketepatan klasifikasi, kemampuan mengenali kelas sebenarnya, maupun keseimbangan antara keduanya.

Confusion matrix

Confusion matrix merupakan alat evaluasi yang menggambarkan perbandingan antara hasil klasifikasi model dan label sebenarnya dari data uji. Matriks ini menampilkan jumlah klasifikasi yang benar maupun salah untuk setiap kelas sentimen. Struktur umum *confusion matrix* ditunjukkan pada tabel berikut.

Tabel 3. 11 Struktur Umum *Confusion matrix*

	Klasifikasi Positif	Klasifikasi Netral	Klasifikasi Negatif
Aktual Positif	<i>True Positif (TP)</i>	<i>False Netral (FN₁)</i>	<i>False Negatif (FN₂)</i>
Aktual Netral	<i>False Positif (FP₁)</i>	<i>True Netral (TN)</i>	<i>False Negatif (FN₃)</i>
Aktual Negatif	<i>False Positif (FP₂)</i>	<i>False Netral (FP₃)</i>	<i>True Negatif (TN₂)</i>

Keterangan :

TP = jumlah data positif yang berhasil diprediksi benar oleh model.

TN = data non positif yang berhasil diprediksi bukan sebagai positif.

FP = jumlah data bukan positif, tetapi salah diprediksi sebagai positif.

FN = jumlah data positif yang salah diprediksi ke kelas lain.

Karena penelitian ini memiliki tiga kelas sentimen (positif, netral, negatif), evaluasi dilakukan secara per kelas (*per-class evaluation*) dan dirata-ratakan untuk menghasilkan nilai *macro average*.

Contoh Perhitungan Manual Evaluasi Model

Untuk memberikan ilustrasi mengenai pengukuran kinerja model, berikut contoh hasil klasifikasi ilustratif dari model *Logistic Regression* terhadap 15 data uji. Data ini digunakan untuk memperlihatkan cara penghitungan metrik evaluasi secara manual.

Tabel 3. 12 Contoh *Confusion matrix* Hasil Klasifikasi *Logistic Regression* (Ilustratif)

Aktual / Klasifikasi	Positif	Netral	Negatif	Total Aktual
Positif	4	1	0	5
Netral	1	3	1	5
Negatif	0	1	4	5
Total Klasifikasi	5	5	5	15

Berdasarkan tabel tersebut:

1. Untuk kelas Positif: TP = 4, FP = 1, FN = 1
2. Untuk kelas Netral: TP = 3, FP = 2, FN = 2
3. Untuk kelas Negatif: TP = 4, FP = 1, FN = 1

Nilai-nilai ini kemudian digunakan untuk menghitung metrik evaluasi menggunakan rumus yang telah dijelaskan pada Bab II (Landasan Teori). Perhitungan metrik dilakukan per kelas, kemudian dirata-ratakan secara makro (*macro average*) agar hasil evaluasi lebih proporsional terhadap jumlah kelas. Berikut contoh perhitungannya:

Kelas Positif :

$$\begin{aligned} Precision &= \frac{4}{4+1} = 0,8, & Recall &= \frac{4}{4+1} = 0,8 \\ F1 &= 2 \times \frac{0,8 \times 0,8}{0,8 + 0,8} = 0,8 \end{aligned} \quad (3.8)$$

Kelas Netral :

$$\begin{aligned} Precision &= \frac{3}{3+2} = 0,6, & Recall &= \frac{3}{3+2} = 0,6 \\ F1 &= 2 \times \frac{0,6 \times 0,6}{0,6 + 0,6} = 0,6 \end{aligned} \quad (3.9)$$

Kelas Negatif :

$$\begin{aligned} Precision &= \frac{4}{4+1} = 0,8, & Recall &= \frac{4}{4+1} = 0,8 \\ F1 &= 2 \times \frac{0,8 \times 0,8}{0,8 + 0,8} = 0,8 \end{aligned} \quad (3.10)$$

Kemudian nilai rata-rata (*macro average*) dihitung :

$$Accuracy = \frac{4 + 3 + 4}{15} = \frac{11}{15} = 0,733 \quad (3.11)$$

$$Precision_{avg} = \frac{0,8 + 0,6 + 0,8}{3} = 0,733 \quad (3.12)$$

$$Recall_{avg} = \frac{0,8 + 0,6 + 0,8}{3} = 0,733 \quad (3.13)$$

$$F1_{avg} = \frac{0,8 + 0,6 + 0,8}{3} = 0,733 \quad (3.14)$$

Hasil evaluasi tersebut menunjukkan bahwa model *Logistic Regression* mampu mengklasifikasikan sentimen dengan tingkat akurasi sekitar 73,3%, dan performa antar kelas relatif seimbang.

Tabel 3. 13 Hasil Perhitungan Metrik Evaluasi (Contoh Manual)

Kelas Sentimen	<i>Precision</i>	<i>Recall</i>	F1-score
Positif	0,80	0,80	0,80
Netral	0,60	0,60	0,60
Negatif	0,80	0,80	0,80
Rata-Rata (<i>Macro-Average</i>)	0,73	0,73	0,73

Tahap evaluasi ini berfungsi untuk memastikan bahwa model tidak hanya memiliki akurasi yang tinggi, tetapi juga dapat mengklasifikasikan setiap kelas sentimen dengan proporsional. Metrik-metrik tersebut memberikan dasar yang kuat untuk menilai keandalan sistem dalam mengenali emosi atau opini pengguna dari ulasan produk *Skincare*.

3.1.6 Output : Hasil Evaluasi

Tahapan analisis kesalahan (error analysis) dilakukan untuk meninjau hasil klasifikasi yang tidak sesuai antara label sebenarnya dan label hasil klasifikasi model.

Tujuan utama dari tahap ini adalah untuk memahami penyebab kesalahan klasifikasi dan mengidentifikasi pola atau karakteristik teks yang sulit dikenali oleh model *Logistic Regression*.

Analisis kesalahan menjadi bagian penting dalam penelitian berbasis pembelajaran mesin, karena dapat memberikan wawasan mengenai keterbatasan model serta menjadi dasar bagi peningkatan akurasi pada penelitian selanjutnya.

Identifikasi Kesalahan Klasifikasi

Setelah proses evaluasi kinerja model dilakukan, sebagian data uji menghasilkan klasifikasi yang tidak sesuai dengan label aslinya. Kesalahan ini dapat terjadi karena berbagai faktor, seperti kompleksitas bahasa, konteks ganda, atau penggunaan kata yang ambigu.

Berikut contoh data ilustratif hasil klasifikasi yang salah diklasifikasikan oleh model *Logistic Regression*.

Tabel 3. 14 Contoh Data Salah Klasifikasi (Ilustratif)

No	Ulasan Asli	Label Aktual	Klasifikasi Model	Keterangan Kesalahan
1	“Facewash-nya murah dan gampang dibeli di minimarket, tapi bikin kulit jadi kering banget.”	Negatif	Netral	Mengandung dua sentimen (positif “murah”, negatif “kering”) → model bingung karena konteks campuran.
2	“Bagus sih, tapi di aku nggak terlalu ngefek.”	Netral	Positif	Kata “bagus” memberi bobot positif tinggi, meskipun ada frasa penolakan di akhir.
3	“Biasa aja, ga buruk tapi juga ga bagus review orang.”	Netral	Negatif	Model cenderung menangkap kata “buruk” tanpa memahami struktur pembandingan “tidak seburuk”.
4	“Scarlet ini wanginya enak banget, tapi jerawat malah nambah.”	Negatif	Positif	Kata “enak banget” menimbulkan bobot positif lebih besar dari kata “jerawat” yang seharusnya mengindikasikan negatif.
5	“Lumayan sih untuk harga segini, cuma agak bikin perih kalau sering dipakai.”	Negatif	Netral	Model kesulitan mengenali intensitas kata “agak” yang melemahkan makna negatif.

Tabel di atas bersifat ilustratif untuk menunjukkan contoh pola kesalahan klasifikasi yang sering muncul dalam analisis sentimen teks. Data aktual hasil

klasifikasi model ditampilkan dan dibahas secara rinci pada Bab IV (Hasil dan Pembahasan).

3.1.7 *Pseudocode* Proses Klasifikasi Sentimen Menggunakan *Logistic Regression*

Pseudocode berikut disusun untuk menggambarkan alur logika proses klasifikasi sentimen pada ulasan produk *Skincare* menggunakan algoritma *Logistic Regression*. *Pseudocode* ini mencakup tahapan pembagian data, representasi fitur menggunakan TF-IDF, proses pelatihan model, klasifikasi data uji, serta evaluasi performa model menggunakan metrik klasifikasi dan *confusion matrix*.

```
Function TrainAndEvaluateLogReg(dataset, split_ratio, C_value)
Input:
    dataset          → data ulasan (stem_reviews) dan label sentimen
    split_ratio      → rasio data uji (0.2 atau 0.3)
    C_value          → nilai parameter C (1.0 atau 10)
Output:
    model            → model Logistic Regression terlatih
    evaluation_metrics → hasil evaluasi performa
    confusion_matrix  → matriks kesalahan klasifikasi
Step 1: Pisahkan fitur dan label
    reviews ← dataset['stem_reviews']
    labels ← dataset['sentimen'] (sudah dalam bentuk numerik)
Step 2: Lakukan pembagian data
    (X_train, X_test, y_train, y_test) ← TrainTestSplit(
        reviews, labels,
        test_size = split_ratio,
        random_state = 42,
        stratify = labels
    )
Step 3: Lakukan vektorisasi menggunakan TF-IDF
    vectorizer ← TFIDF(max_features = 5000, ngram_range = (1,2))
    X_train_vec ← vectorizer.FitTransform(X_train)
    X_test_vec ← vectorizer.Transform(X_test)
Step 4: Inisialisasi model Logistic Regression
    model ← LogisticRegression(
        solver = "lbfgs",
        penalty = "l2",
        C = C_value,
        max_iter = 2000
    )
Step 5: Latih model
    model.Fit(X_train_vec, y_train)
Step 6: Klasifikasi data uji
```

```

y_pred ← model.Predict(X_test_vec)
Step 7: Hitung evaluasi performa
evaluation_metrics ← ClassificationReport(y_test, y_pred)
confusion_matrix ← ConfusionMatrix(y_test, y_pred)
Step 8: Kembalikan semua hasil
Return (model, evaluation_metrics, confusion_matrix)
End Function

```

3.2 Skenario Pengujian

Skenario pengujian pada penelitian ini dirancang untuk mengevaluasi kinerja algoritma *Logistic Regression* dalam mengklasifikasikan sentimen ulasan produk *Skincare* ke dalam tiga kategori, yaitu positif, negatif, dan netral. Pengujian dilakukan secara sistematis melalui pembagian dataset menjadi data latih dan data uji menggunakan fungsi *train_test_split*. Setiap skenario dijalankan secara independen agar pengaruh variasi parameter dan rasio pembagian data dapat diamati secara empiris.

Parameter C pada *Logistic Regression* mengatur kekuatan regularisasi. Nilai C yang lebih besar mengurangi regularisasi sehingga model menjadi lebih kompleks, sedangkan nilai C kecil meningkatkan regularisasi untuk mencegah *overfitting*. Variasi parameter ini dipilih agar penelitian dapat mengamati dampak regularisasi terhadap model sentimen berbasis TF-IDF.

Pengujian dilakukan menggunakan dua rasio pembagian data yang umum digunakan dalam penelitian analisis sentimen, yaitu 80:20 dan 70:30. Rasio 80:20 digunakan sebagai skenario utama untuk memaksimalkan data latih, sedangkan rasio 70:30 digunakan untuk melihat pengaruh pengurangan data latih terhadap performa model. Kombinasi dua rasio tersebut dengan dua nilai parameter C menghasilkan empat skenario pengujian yang ditunjukkan pada Tabel 3.16.

Evaluasi performa model pada setiap skenario menggunakan metrik akurasi, *precision*, *recall*, *F1-score*, dan *confusion matrix*. Kelima metrik ini dipilih karena mampu menggambarkan kinerja klasifikasi secara menyeluruh pada ketiga kelas sentimen.

Tabel 3. 15 Skenario Pengujian

Skenario /Model	Metode Split Data	Rasio Split	Parameter Model	Output yang Diuji	Keterangan/Hasil
1	Random Split	80 : 20	Default <i>Logistic Regression</i> (solver='lbfgs', penalty='l2', C=1.0, max_iter=2000)	<i>Accuracy</i> , <i>Precision</i> , <i>Recall</i> , <i>F1-Score</i> , <i>Confusion matrix</i>	Skenario utama menggunakan parameter default pada rasio data umum 80:20.
2	Random Split	80 : 20	<i>Logistic Regression</i> (solver='lbfgs', penalty='l2', C=10, max_iter=2000)	<i>Accuracy</i> , <i>Precision</i> , <i>Recall</i> , <i>F1-Score</i> , <i>Confusion matrix</i>	Skenario untuk melihat efek pengurangan regularisasi (tuning C) pada rasio 80:20.
3	Random Split	70 : 30	Default <i>Logistic Regression</i> (solver='lbfgs', penalty='l2', C=1.0, max_iter=2000)	<i>Accuracy</i> , <i>Precision</i> , <i>Recall</i> , <i>F1-Score</i> , <i>Confusion matrix</i>	Mengetahui pengaruh pengurangan data latih terhadap performa model..
4	Random Split	70 : 30	<i>Logistic Regression</i> (solver='lbfgs', penalty='l2', C=10, max_iter=2000)	<i>Accuracy</i> , <i>Precision</i> , <i>Recall</i> , <i>F1-Score</i> , <i>Confusion matrix</i>	Kombinasi tuning parameter C pada rasio data berbeda untuk melihat konsistensi efeknya.

Melalui keempat skenario tersebut, penelitian ini dapat memberikan gambaran mengenai stabilitas dan efektivitas *Logistic Regression* dalam analisis sentimen ulasan *Skincare* pada berbagai kondisi pembagian data dan tingkat regularisasi model. Perbandingan antar skenario akan dianalisis pada Bab IV.

BAB IV

HASIL DAN PEMBAHASAN

4.1 Deskripsi Umum Pengujian

Bab ini menyajikan hasil pemodelan dan evaluasi kinerja model *Logistic Regression* dalam mengklasifikasikan sentimen ulasan produk *Skincare* di *E-Commerce* ke dalam tiga kelas, yaitu positif, negatif, dan netral.

Pengujian model dilakukan menggunakan empat skenario yang berbeda dengan variasi parameter regularisasi ($C = 1$ dan $C = 10$) serta rasio pembagian data latih dan data uji (80:20 dan 70:30). Variasi skenario ini digunakan untuk mengamati pengaruh konfigurasi parameter terhadap kinerja model *Logistic Regression*.

Tahapan representasi fitur menggunakan TF-IDF dan pembagian data merupakan bagian dari proses pemodelan yang telah dijelaskan pada Bab III. Oleh karena itu, pada bab ini hasil penelitian difokuskan pada keluaran model *Logistic Regression* berupa klasifikasi sentimen serta evaluasi kinerjanya yang diukur menggunakan *confusion matrix* dan metrik *accuracy*, *precision*, *recall*, dan *F1-score*.

4.2 Hasil Pemodelan *Logistic Regression*

Hasil pemodelan *Logistic Regression* pada penelitian ini berupa klasifikasi label sentimen dan nilai probabilitas kelas untuk setiap ulasan pada data uji. Model menentukan kelas sentimen berdasarkan probabilitas tertinggi yang dihasilkan dari proses klasifikasi.

Karena jumlah data uji yang besar, hasil klasifikasi individual tidak disajikan secara keseluruhan. Oleh karena itu, Tabel 4.1 di bawah ini menampilkan hasil klasifikasi model *Logistic Regression* pada sepuluh ulasan, yang terdiri dari lima ulasan awal dan lima ulasan akhir, sebagai ilustrasi keluaran model sebelum dilakukan evaluasi kinerja secara kuantitatif. Hasil klasifikasi yang disajikan berasal dari skenario C10_80_20, yaitu skenario dengan konfigurasi parameter yang menghasilkan kinerja terbaik berdasarkan hasil evaluasi.

Tabel 4. 1 Hasil Klasifikasi Model *Logistic Regression*

no	Stem Reviews	Label Asli	Label Klasifikasi	Probabilitas Negatif	Probabilitas Netral	Probabilitas Positif
1	produk give away soco box terima kasih sociolla kayak cocok rambut mudah rontok kadang ketombe ingredients sesuai butuh rambut kulit kepala dapat bareng conditionernya moga cocok keluarga	positif	positif	0,008	0,019	0,973
2	tekstur lengket aplikasi muka warna beda ulas kuning coklat kya campur	negatif	negatif	0,641	0,313	0,046
3	kulit kering cocok pakai ngelembabkan kulit bau nivea night nourish lotion wangi xd mudah serap kulit cerah malam pakai atur	netral	positif	0,014	0,031	0,955
4	tahan produk rambut kering kusut lembut efek tahan harga murah kualitas	negatif	negatif	0,926	0,029	0,046
5	verifikasi identitas seller platform	netral	netral	0,099	0,815	0,086
...
1195	coba lumayan wajah bersih lumayan kering wajah tipe oily bilang kayak pilih alih produk kulit kering	netral	netral	0,023	0,958	0,019
1196	lembap hari bersih wajah jaga kelembaban kulit	netral	netral	0,004	0,971	0,025
1197	paket	netral	netral	0,028	0,801	0,171

Berdasarkan Tabel 4.1, dapat dilihat bahwa model *Logistic Regression* menghasilkan klasifikasi sentimen dan nilai probabilitas kelas untuk setiap ulasan. Label sentimen ditentukan berdasarkan nilai probabilitas tertinggi yang dihasilkan

oleh model. Penilaian evaluasi kinerja model dilakukan pada subbab berikutnya melalui *confusion matrix* dan metrik evaluasi.

4.3 Hasil Evaluasi

Subbab ini menyajikan hasil evaluasi kinerja model *Logistic Regression* dalam mengklasifikasikan sentimen ulasan produk *Skincare* ke dalam tiga kelas, yaitu positif, netral, dan negatif. Evaluasi dilakukan pada data uji untuk menilai kualitas klasifikasi model menggunakan *confusion matrix* serta metrik *accuracy*, *precision*, *recall*, dan *F1-score* dengan pendekatan rata-rata makro (*macro average*).

Pengujian dilakukan pada empat skenario dengan variasi nilai parameter C (1 dan 10) serta rasio pembagian data latih dan data uji (80:20 dan 70:30). Hasil evaluasi tiap skenario disajikan secara terpisah untuk melihat pengaruh konfigurasi parameter terhadap performa model, kemudian dibandingkan guna menentukan skenario dengan kinerja terbaik.

4.3.1 Skenario1

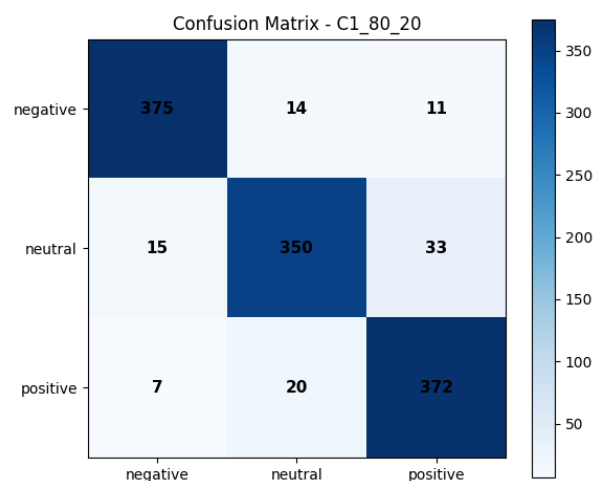
Tabel 4.2 menyajikan hasil evaluasi kinerja model *Logistic Regression* pada skenario C = 1 dengan rasio pembagian data latih dan data uji sebesar 80:20. Berdasarkan tabel tersebut, model memperoleh nilai *accuracy* sebesar 0,916, *precision (macro)* sebesar 0,917, *recall (macro)* sebesar 0,916, serta *F1-score (macro)* sebesar 0,916. Nilai-nilai ini menunjukkan bahwa model mampu melakukan klasifikasi sentimen dengan tingkat ketepatan dan keseimbangan yang

baik pada seluruh kelas, sehingga konfigurasi parameter $C = 1.0$ dengan proporsi data latih yang lebih besar memberikan performa yang stabil.

Tabel 4. 2 Hasil Evaluasi $C=1$ Split 80:20

Skenario	<i>Accuracy</i>	<i>Precision (macro)</i>	<i>Recall (macro)</i>	<i>F1-score (macro)</i>
$C=1$ – Split 80:20	0,9164578111	0,9167578284	0,9164092706	0,9163509862

Gambar 4.1 memperlihatkan *confusion matrix* dari skenario C1_80_20 yang menggambarkan distribusi hasil klasifikasi model terhadap masing-masing kelas sentimen. Terlihat bahwa sebagian besar data pada kelas negatif, netral, dan positif berhasil diklasifikasikan dengan benar. Kesalahan klasifikasi relatif kecil dan umumnya terjadi pada kelas netral yang diklasifikasikan sebagai positif. Pola ini menunjukkan bahwa model masih mengalami ambiguitas pada kelas netral, namun secara keseluruhan kemampuan klasifikasi tetap tergolong baik.



Gambar 4. 1 *Confusion matrix* $C=1$ Split 80:20

4.3.2 Skenario2

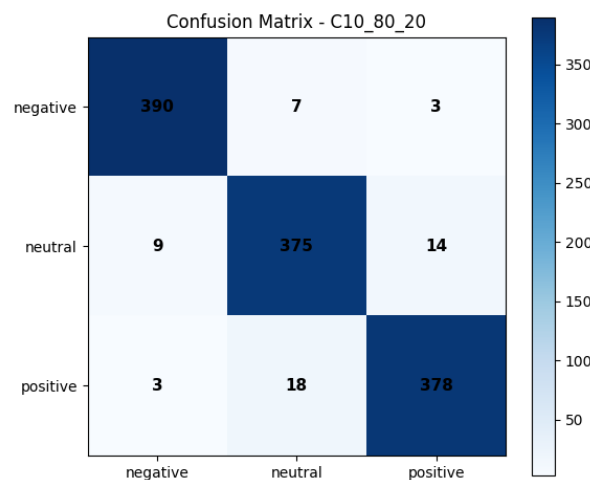
Tabel 4.3 menunjukkan hasil evaluasi model *Logistic Regression* pada skenario $C = 10$ dengan rasio pembagian data 80:20. Pada konfigurasi ini, model

memperoleh nilai *accuracy*, *precision (macro)*, *recall (macro)*, dan *F1-score (macro)* yang sama-sama berada pada angka 0,955. Peningkatan nilai *C* menyebabkan regularisasi menjadi lebih longgar sehingga model memiliki fleksibilitas lebih tinggi dalam menyesuaikan parameter. Hal ini berdampak pada peningkatan performa secara konsisten pada seluruh metrik evaluasi dibandingkan skenario sebelumnya.

Tabel 4. 3 Hasil Evaluasi C=10_Split 80:20

Skenario	<i>Accuracy</i>	<i>Precision (macro)</i>	<i>Recall (macro)</i>	<i>F1-score (macro)</i>
C=10 – Split 80:20	0,9548872180	0,9548704263	0,9548598254	0,9548530868

Gambar 4.2 menampilkan *confusion matrix* pada skenario C10_80_20. Matriks ini menunjukkan bahwa kesalahan klasifikasi pada ketiga kelas sentimen sangat rendah, dengan dominasi klasifikasi yang benar pada seluruh kelas. Distribusi hasil klasifikasi yang merata menandakan bahwa model mampu mengenali karakteristik masing-masing sentimen secara optimal. Oleh karena itu, skenario ini dinilai sebagai konfigurasi terbaik dalam penelitian ini.



Gambar 4. 2 *Confusion matrix* C=10_Split 80:20

Pada tabel 4.4 menampilkan hasil evaluasi per kelas pada skenario C10_80_20 menunjukkan performa model yang sangat baik dan stabil pada ketiga kategori sentimen. Kelas negatif memperoleh nilai *precision* sebesar 0,970 dan *recall* sebesar 0,975, yang menunjukkan bahwa model dapat mengidentifikasi ulasan negatif dengan tingkat kesalahan yang sangat rendah. Kelas netral memiliki nilai *precision* sebesar 0,938 dan *recall* sebesar 0,942; meskipun sedikit lebih rendah dibanding kelas lainnya, performa pada kelas ini tetap berada pada kategori sangat baik. Sementara itu, kelas positif memperoleh *precision* sebesar 0,957 dan *recall* sebesar 0,947, yang menandakan keseimbangan deteksi dan akurasi klasifikasi pada kategori positif. Nilai rata-rata makro (*macro average*) untuk *precision*, *recall*, dan *F1-score* berada pada angka 0,955, yang sekaligus merupakan nilai tertinggi di antara seluruh skenario. Hasil ini memperkuat temuan bahwa konfigurasi parameter $C = 10$ dengan pembagian data 80:20 menghasilkan performa paling optimal dalam mengklasifikasikan ketiga kelas sentimen secara konsisten.

Tabel 4. 4 Hasil Evaluasi setiap Kelas C=10 Split 80:20

Kelas Sentimen	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>	Jumlah Data
Negatif	0,970	0,975	0,973	400
Netral	0,938	0,942	0,940	398
Positif	0,957	0,947	0,952	399
Rata-rata (<i>macro</i>)	0,955	0,955	0,955	

4.3.3 Skenario3

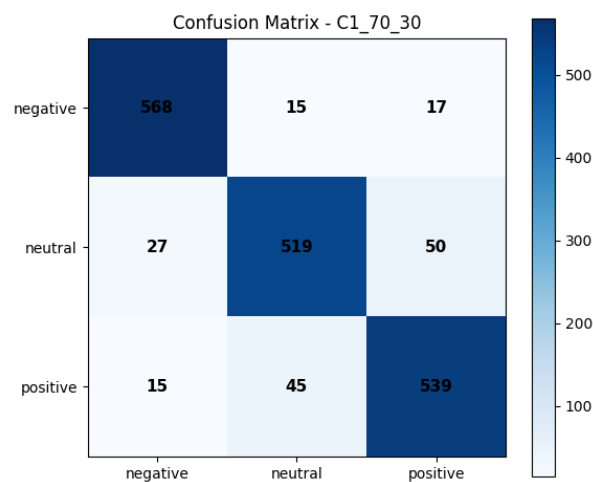
Tabel 4.5 menyajikan hasil evaluasi model pada skenario $C = 1$ dengan rasio pembagian data 70:30. Berdasarkan tabel tersebut, model memperoleh *accuracy* sebesar 0,906 dan *F1-score (macro)* sebesar 0,906. Nilai ini sedikit lebih rendah dibandingkan skenario dengan rasio 80:20, yang menunjukkan bahwa pengurangan

jumlah data latih berpengaruh terhadap kemampuan model dalam mempelajari pola data secara optimal.

Tabel 4. 5 Hasil Evaluasi C=1 Split 70:30

Skenario	<i>Accuracy</i>	<i>Precision (macro)</i>	<i>Recall (macro)</i>	<i>F1-score (macro)</i>
C=1 – Split 70:30	0,9058495821	0,9056531806	0,9057683636	0,9056176798

Gambar 4.3 memperlihatkan *confusion matrix* pada skenario C1_70_30. Terlihat bahwa kesalahan klasifikasi lebih banyak terjadi dibandingkan skenario sebelumnya, terutama pada kelas netral yang sering diklasifikasikan sebagai positif. Hal ini mengindikasikan bahwa dengan jumlah data latih yang lebih sedikit, model mengalami kesulitan dalam membedakan kelas sentimen yang bersifat ambigu.



Gambar 4. 3 *Confusion matrix* C=1_Split 70:30

4.3.4 Skenario4

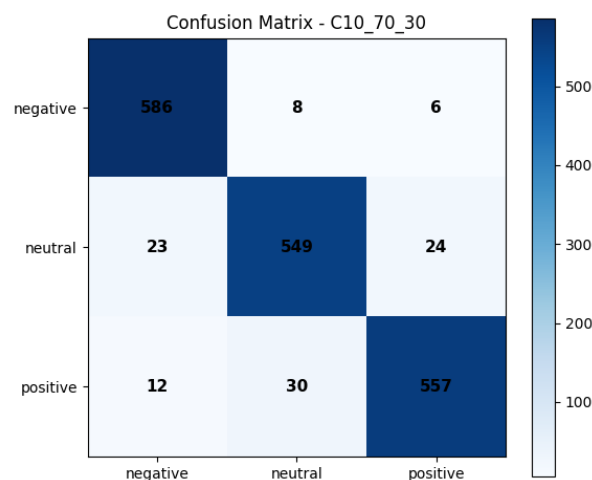
Tabel 4.6 menunjukkan hasil evaluasi model pada skenario C = 10 dengan rasio pembagian data 70:30. Model memperoleh nilai *accuracy* sebesar 0,943 dan *F1-score (macro)* sebesar 0,942. Peningkatan nilai C terbukti mampu memperbaiki

performa model dibandingkan skenario $C = 1$ pada rasio yang sama, meskipun jumlah data latih lebih sedikit.

Tabel 4. 6 Hasil Evaluasi $C=10$ Split 70:30

Skenario	Accuracy	Precision (macro)	Recall (macro)	F1-score (macro)
C=10 – Split 70:30	0,9426183844	0,9425986735	0,9425635816	0,942436490

Gambar 4.4 menampilkan *confusion matrix* skenario C10_70_30 yang menunjukkan distribusi klasifikasi yang lebih akurat dibandingkan skenario C1_70_30. Kesalahan klasifikasi pada ketiga kelas relatif berkurang, yang menandakan bahwa nilai C yang lebih besar membantu model dalam mengenali fitur penting meskipun data latih terbatas.



Gambar 4. 4 *Confusion matrix* $C=10$ _Split 70:30

4.3.5 Perbandingan Antar Skenario

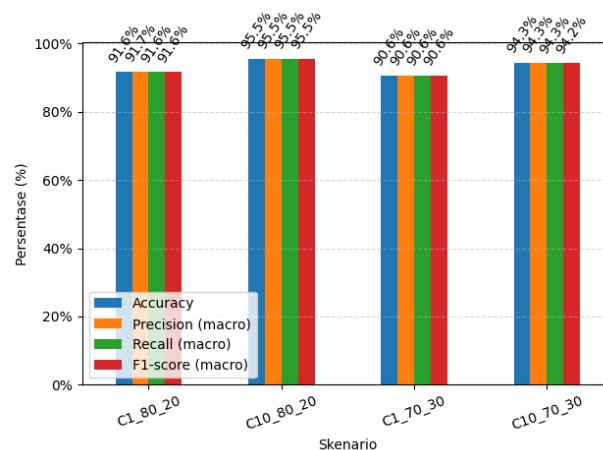
Tabel 4.7 menyajikan perbandingan kinerja model *Logistic Regression* pada keempat skenario pengujian. Berdasarkan tabel tersebut, terlihat bahwa skenario C10_80_20 menghasilkan nilai tertinggi pada seluruh metrik evaluasi, diikuti oleh

skenario C10_70_30. Sementara itu, skenario dengan nilai $C = 1$ menunjukkan performa yang lebih rendah, terutama pada rasio pembagian data 70:30. Hal ini menegaskan bahwa nilai parameter C dan jumlah data latih memiliki pengaruh signifikan terhadap kinerja model.

Tabel 4. 7 Hasil Evaluasi Perbandingan 4Skenario

Skenario	<i>Accuracy</i>	<i>Precision (macro)</i>	<i>Recall (macro)</i>	<i>F1-score (macro)</i>
C=1 – Split 80:20	0,916	0,917	0,916	0,916
C=10 – Split 80:20	0,955	0,955	0,955	0,955
C=1 – Split 70:30	0,906	0,906	0,906	0,906
C=10 – Split 70:30	0,943	0,943	0,943	0,942

Gambar 4.5 menampilkan grafik perbandingan *accuracy*, *precision*, *recall*, dan *F1-score* antar skenario. Grafik tersebut memperlihatkan tren peningkatan performa ketika nilai C dinaikkan dari 1 menjadi 10 serta ketika proporsi data latih diperbesar. Dengan demikian, dapat disimpulkan bahwa kombinasi nilai C yang lebih besar dan rasio data latih yang lebih tinggi memberikan hasil evaluasi terbaik pada penelitian ini.



Gambar 4. 5 Grafik Perbandingan Akurasi, *Precision*, *Recall*, dan *F1-score* Antar Skenario

Hasil pada Tabel dan Gambar diatas menunjukkan bahwa performa model *Logistic Regression* dipengaruhi oleh dua faktor utama, yaitu nilai parameter C dan rasio pembagian data.

1. Pengaruh Nilai Parameter C

Nilai $C = 10$ consistently menghasilkan performa lebih tinggi dibandingkan $C = 1.0$ pada kedua rasio pembagian data. Pada rasio 80:20. *F1-score* meningkat dari 0,9163 menjadi 0,9548. Kenaikan sebesar +3,85%. Sedangkan pada rasio 70:30. *F1-score* meningkat dari 0,9056 menjadi 0,9424. Kenaikan sebesar +3,68%. Peningkatan ini menunjukkan bahwa nilai C yang lebih besar membuat regularisasi lebih longgar, sehingga model dapat belajar pola teks secara lebih fleksibel tanpa dibatasi terlalu kuat oleh penalty L2.

2. Pengaruh Rasio Pembagian Data

Rasio pembagian data 80% data latih dan 20% data uji menghasilkan performa yang lebih baik dibandingkan rasio 70%:30% pada kedua nilai parameter C yang digunakan. Pada $C = 1.0$, peningkatan rasio data latih menyebabkan *F1-score* naik dari 0,9056 menjadi 0,9163, yang menunjukkan perbaikan keseimbangan antara *precision* dan *recall*. Sementara itu, pada $C = 10$, nilai *F1-score* juga mengalami peningkatan yang signifikan dari 0,9424 menjadi 0,9548. Hasil ini mengindikasikan bahwa penambahan jumlah data latih memberikan kontribusi positif terhadap kemampuan model dalam mempelajari pola sentimen. Dengan data latih yang lebih banyak, model mampu melakukan generalisasi dengan lebih baik, khususnya dalam menangkap variasi struktur dan konteks kalimat pada ulasan produk *Skincare*.

3. Skenario Terbaik

Skenario C10_80_20 merupakan konfigurasi terbaik secara keseluruhan karena menunjukkan kinerja paling optimal pada seluruh metrik evaluasi yang digunakan. Nilai *accuracy* pada skenario ini mencapai 0,9548, yang menandakan tingkat ketepatan klasifikasi model yang sangat tinggi. Selain itu, *precision* sebesar 0,9548 menunjukkan bahwa model mampu menghasilkan klasifikasi sentimen yang akurat dengan tingkat kesalahan yang minimal. Nilai *recall* yang juga mencapai 0,9548 mengindikasikan kemampuan model dalam mengenali seluruh kelas sentimen secara konsisten. Selanjutnya, F1-score sebesar 0,9548 mencerminkan keseimbangan yang sangat baik antara *precision* dan *recall*. Keunggulan skenario ini terlihat jelas pada grafik perbandingan, di mana batang nilai untuk seluruh metrik berada pada posisi paling tinggi dibandingkan skenario lainnya. Kinerja unggul tersebut dipengaruhi oleh penggunaan nilai $C = 10$ yang memberikan fleksibilitas regularisasi lebih optimal sehingga model dapat mempelajari bobot kata secara lebih efektif. Selain itu, rasio data latih 80% memungkinkan model memperoleh lebih banyak sampel pelatihan, sehingga mampu menangkap pola linguistik dalam ulasan produk secara lebih komprehensif.

4. Kualitas Konsistensi Antar Metrik

Seluruh metrik evaluasi (*accuracy*, *precision*, *recall*, dan F1-score) menunjukkan pola yang konsisten pada setiap skenario. Skenario dengan akurasi tertinggi selalu memiliki nilai F1-score tertinggi, sehingga tidak terlihat adanya anomali antar metrik. Konsistensi ini menunjukkan bahwa model mampu melakukan klasifikasi secara seimbang pada seluruh kelas sentimen dan tidak

menunjukkan kecenderungan bias terhadap kelas tertentu. Dengan demikian, *Logistic Regression* dapat dikatakan bekerja secara stabil pada dataset yang memiliki distribusi kelas yang seimbang.

Berdasarkan keseluruhan hasil, dapat disimpulkan bahwa Parameter $C = 10$ lebih efektif dibandingkan $C = 1.0$. Rasio 80:20 menghasilkan performa lebih tinggi dibandingkan 70:30. Skenario C10_80_20 adalah konfigurasi terbaik dan paling direkomendasikan untuk dataset ulasan *Skincare* ini.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Penelitian ini bertujuan untuk menjawab pertanyaan mengenai bagaimana penerapan TF-IDF dan *Logistic Regression* dalam mengklasifikasikan sentimen ulasan produk *Skincare* pada platform *E-Commerce*. Berdasarkan tahapan yang dilakukan—meliputi pemuatan dataset publik, representasi teks menggunakan TF-IDF, serta pelatihan model *Logistic Regression* pada empat skenario pengujian—dapat disimpulkan bahwa metode tersebut mampu bekerja secara efektif dalam memetakan ulasan ke dalam tiga kategori sentimen, yaitu positif, netral, dan negatif.

Hasil evaluasi menunjukkan bahwa seluruh skenario menghasilkan performa yang tinggi dengan nilai akurasi, *precision*, *recall*, dan *F1-score (macro)* berada di atas 90%. Model dengan kinerja terbaik diperoleh pada skenario C10_80_20, yaitu *Logistic Regression* dengan parameter $C = 10$ dan pembagian data latih-uji 80:20, yang mencapai akurasi tertinggi dibandingkan skenario lainnya. Temuan ini mengindikasikan bahwa penyesuaian parameter regularisasi (C) serta proporsi data latih memengaruhi kualitas klasifikasi, dan tuning parameter memberikan peningkatan performa yang signifikan dibandingkan penggunaan parameter default.

Secara keseluruhan, penelitian ini membuktikan bahwa kombinasi TF-IDF dan *Logistic Regression* merupakan pendekatan yang efisien dan akurat untuk analisis sentimen ulasan produk *Skincare* berbahasa Indonesia, sehingga dapat

menjadi dasar bagi pengembangan sistem analisis sentimen yang lebih lanjut pada domain *E-Commerce*.

5.2 Saran

Berdasarkan temuan penelitian ini, beberapa saran yang dapat diberikan untuk penelitian selanjutnya adalah sebagai berikut.

1. Mengembangkan model dengan algoritma pembandingan.

Penelitian berikutnya dapat mencoba algoritma lain seperti SVM, *Naive Bayes*, Random Forest, atau *deep learning* (misalnya LSTM, BERT) untuk memperoleh perbandingan performa yang lebih luas.

2. Menggunakan dataset yang lebih besar dan lebih beragam.

Penelitian selanjutnya dapat memperluas cakupan ulasan dengan menambahkan data dari berbagai kategori produk atau platform *E-Commerce* lainnya untuk meningkatkan generalisasi model.

3. Eksplorasi metode optimasi model.

Misalnya melakukan *hyperparameter tuning* secara lebih sistematis menggunakan Grid Search atau Random Search untuk mendapatkan konfigurasi parameter yang lebih optimal.

4. Menganalisis struktur bahasa dan konteks yang lebih kompleks. Pendekatan NLP modern seperti *word embedding* (Word2Vec, FastText) atau transformer-based embedding dapat diteliti lebih lanjut untuk menangkap konteks semantik yang tidak terwakili oleh TF-IDF.

5. Mengimplementasikan model dalam bentuk aplikasi sederhana.

Misalnya dashboard analisis sentimen berbasis web untuk memantau ulasan konsumen secara otomatis sehingga dapat membantu pelaku bisnis dalam pengambilan keputusan.

DAFTAR PUSTAKA

- Al-Qur'an. (n.d.). Surah Al-Hujurat [49:6].
- Amalia, A., Putra, A. R., & Prasetyo, E. (2019). Analisis sentimen ulasan produk menggunakan metode *Logistic Regression*. *Jurnal Ilmiah Komputer dan Informatika (JIKI)*, 8(1), 34–41. <https://doi.org/10.33365/jiki.v8i1.613>
- DataReportal. (2023). Digital 2023: Indonesia. <https://datareportal.com/reports/digital-2023-indonesia>
- Fitriani, R., Syahrizal, & Hanifah, N. (2020). Perbandingan pendekatan lexicon-based dan *machine learning* untuk analisis sentimen. *Jurnal Teknologi dan Sistem Komputer*, 8(3), 249–257. <https://doi.org/10.14710/jtsiskom.8.3.2020.249-257>
- Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied Logistic Regression* (3rd ed.). Wiley. <https://doi.org/10.1002/9781118548387>
- Kementerian Perdagangan Republik Indonesia. (2023). Laporan *E-Commerce Indonesia 2023*. <https://www.kemendag.go.id/id/E-Commerce2023>
- Kusumaningrum, E., & Nugroho, A. S. (2021). *Sentiment analysis using Logistic Regression and TF-IDF on consumer review*. *Indonesian Journal of Artificial Intelligence and Data Mining*, 4(1), 1–8. <https://doi.org/10.24014/ijaidm.v4i1.13581>
- Liu, B. (2020). *Sentiment analysis: Mining opinions, sentiments, and emotions*. Cambridge University Press. <https://doi.org/10.1017/9781108953987>
- Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to Information Retrieval*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511809071>
- Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, 5(4), 1093–1113. <https://doi.org/10.1016/j.asej.2014.04.011>
- Mufidah, L. (2021). Etika bisnis Islam dalam konteks ekonomi digital. *Jurnal Ekonomi Syariah Indonesia*, 11(2), 145–155. <https://doi.org/10.5281/zenodo.5157320>
- Najmi, E., Masyhuda, L., & Arliza, N. (2025). Sentiment analysis on traveloka hotel reviews using *Logistic Regression and Support Vector Machine with SMOTE imbalance handling*. *BRIGHT JADS (Journal of Applied Data Sciences)*, 6(2), 114–120. <https://doi.org/10.47738/jads.v6i2.637>

- Nugraha, A., Prasetyo, D. E., & Rahmawati, F. (2022). Analisis pengaruh ulasan terhadap strategi pemasaran *E-Commerce*. *Jurnal Sistem Informasi dan Informatika (JSII)*, 6(4), 229–240. <https://doi.org/10.30865/jsii.v6i4.4365>
- Özmen, C. G., & Gündüz, S. (2025). Comparison of machine learning models for sentiment analysis of big turkish web-based data. *Applied Sciences*, 15(5), 2297. <https://doi.org/10.3390/app15052297>
- Pratama, B., & Wulandari, E. (2021). Analisis sentimen ulasan konsumen marketplace menggunakan machine learning. *Jurnal Teknologi Informasi dan Komputer (JTIK)*, 7(1), 12–20. <https://doi.org/10.36085/jtik.v7i1.1125>
- Putra, R. A., & Rahmawati, F. (2022). Perkembangan transaksi *E-Commerce* di Indonesia tahun 2022. *Jurnal Ekonomi Digital Indonesia*, 2(3), 88–96. <https://doi.org/10.32520/jedi.v2i3.152>
- Romadhony, A., Al Faraby, S., Rismala, R., Wisesti, U. N., & Arifianto, A. (2024). Sentiment analysis on a large indonesian product review dataset. *Journal of Information Systems Engineering and Business Intelligence*, 10(1), 167–178. <https://doi.org/10.20473/jisebi.10.1.167-178>
- Sari, N. W., & Fadillah, R. (2021). Analisis sentimen ulasan produk kecantikan di shopee menggunakan *Naive Bayes*. *Jurnal Ilmiah Komputer dan Informatika (JIKI)*, 9(1), 15–24. <https://doi.org/10.33365/jiki.v9i1.512>
- Setiawan, A., & Hidayat, A. (2021). Pengaruh e-WOM terhadap keputusan pembelian produk di marketplace. *Jurnal Manajemen dan Pemasaran Jasa*, 14(1), 55–66. <https://doi.org/10.25105/jmpj.v14i1.8854>
- Statista. (2023). Indonesia: Online shopping behaviour and market insights. <https://www.statista.com/topics/7778/online-shopping-in-indonesia>
- Sutarman, R. Siringoringo, D. Arisandi, E. Kurniawan, & E. B. Nababan. (2024). Model Klasifikasi dengan *Logistic Regression* dan *Recursive Feature Elimination* pada Data Tidak Seimbang. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 11(4), 735–742.
- Syafitri, N., Nurjanah, D., & Yulianto, B. (2020). Analisis sentimen review produk kosmetik menggunakan pendekatan text mining. *Jurnal Teknologi dan Sistem Informasi (JTSI)*, 11(3), 223–232. <https://doi.org/10.29207/jtisi.v11i3.3202>
- Wahyuni, T., Nurhayati, H., & Firmansyah, F. (2020). Analisis sentimen produk di Tokopedia menggunakan *Naive Bayes* dan SVM. *Jurnal Ilmiah Komputer dan Informatika (JIKI)*, 8(2), 102–109. <https://doi.org/10.33365/jiki.v8i2.618>
- Zhang, L., & Luo, Y. (2020). *Logistic Regression* in sentiment analysis: Comparison and performance review. *International Journal of Computational Linguistics Research*, 11(4), 97–105. <https://doi.org/10.31838/ijclr.11.4.97>