

**ANALISIS SENTIMEN ULASAN PELANGGAN PADA TOKO
RETAIL SUPER DAZZLE DI *GOOGLE MAPS*
MENGUNAKAN ALGORITMA *NAÏVE BAYES* UNTUK
EVALUASI KUALITAS LAYANAN**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

YOHANES PANGESTU

22.11.5316

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2026

**ANALISIS SENTIMEN ULASAN PELANGGAN PADA TOKO
RETAIL SUPER DAZZLE DI *GOOGLE MAPS*
MENGUNAKAN ALGORITMA *NAÏVE BAYES* UNTUK
EVALUASI KUALITAS LAYANAN**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

YOHANES PANGESTU

22.11.5316

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2026**

HALAMAN PERSETUJUAN

SKRIPSI

ANALISIS SENTIMEN ULASAN PELANGGAN PADA TOKO RETAIL SUPER DAZZLE DI GOOGLE MAPS MENGGUNAKAN ALGORITMA NAÏVE BAYES UNTUK EVALUASI KUALITAS LAYANAN

yang disusun dan diajukan oleh

Yohanes Pangestu

22.11.5316

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 9 Juni 2026

Dosen Pembimbing,



Theopilus Bayu Sasongko, S.Kom., M.Eng

NIK. 190302375

HALAMAN PENGESAHAN

SKRIPSI

**ANALISIS SENTIMEN ULASAN PELANGGAN PADA TOKO RETAIL SUPER
DAZZLE DI GOOGLE MAPS MENGGUNAKAN ALGORITMA NAÏVE BAYES
UNTUK EVALUASI KUALITAS LAYANAN**

yang disusun dan diajukan oleh

Yohanes Pangestu

22.11.5316

Telah dipertahankan di depan Dewan Penguji
pada tanggal 25 Juni 2026

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Ali Mustopa, S.Kom., M.Kom
NIK. 190302192

Ade Pujianto, M.Kom
NIK. 190302494

Theopilus Bayu Sasongko, S.Kom., M.Eng.
NIK. 190302375



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 25 Juni 2026

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusrini, M.Kom.
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Yohanes Pangestu
NIM : 22.11.5316

Menyatakan bahwa Skripsi dengan judul berikut:

**ANALISIS SENTIMEN ULASAN PELANGGAN PADA TOKO RETAIL
SUPER DAZZLE DI GOOGLE MAPS MENGGUNAKAN ALGORITMA
NAÏVE BAYES UNTUK EVALUASI KUALITAS LAYANAN**

Dosen Pembimbing : Theopilus Bayu Sasongko, S.Kom.,M.Eng

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 25 Juni 2026

Yang Menyatakan,



Yohanes Pangestu

KATA PENGANTAR

Puji syukur penulis panjatkan kepada Tuhan Yang Maha Esa atas segala rahmat, karunia, dan penyertaan-Nya sehingga penulis dapat menyelesaikan skripsi berjudul “Analisis Sentimen Ulasan Pelanggan Pada Toko Retail Super Dazzle Menggunakan Algoritma Naïve Bayes Untuk Evaluasi Kualitas Layanan” dengan baik. Skripsi ini disusun sebagai syarat untuk memperoleh gelar Sarjana pada Program Studi Informatika. Dalam proses penyusunan skripsi ini, penulis memperoleh banyak bantuan, dukungan, serta bimbingan dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis menyampaikan ucapan terima kasih kepada :

1. Bapak Theopilus Bayu Sasongko, S.Kom.,M.Eng selaku dosen pembimbing yang telah memberikan arahan, masukan, dan bimbingan selama proses penyusunan skripsi.
2. Bapak Ali Mustopa, S.Kom., M.Kom dan Bapak Ade Pujianto, M.Kom selaku Tim Dosen Penguji yang telah memberikan kritik dan saran yang telah membangun untuk penyempurnaan skripsi ini.
3. Kedua orang tua dan keluarga tercinta yang telah memberikan doa, dukungan, semangat, serta motivasi kepada penulis.
4. Teman-teman dan seluruh pihak yang telah membantu serta memberikan dukungan selama proses penyusunan skripsi.

Penulis menyadari bahwa skripsi ini masih memiliki kekurangan. Oleh karena itu, penulis mengharapkan kritik dan saran yang membangun demi penyempurnaan penelitian dimasa mendatang. Semoga skripsi ini dapat memberikan manfaat bagi pembaca dan menjadi referensi bagi penelitian selanjutnya.

Yogyakarta, 25 Juni 2026

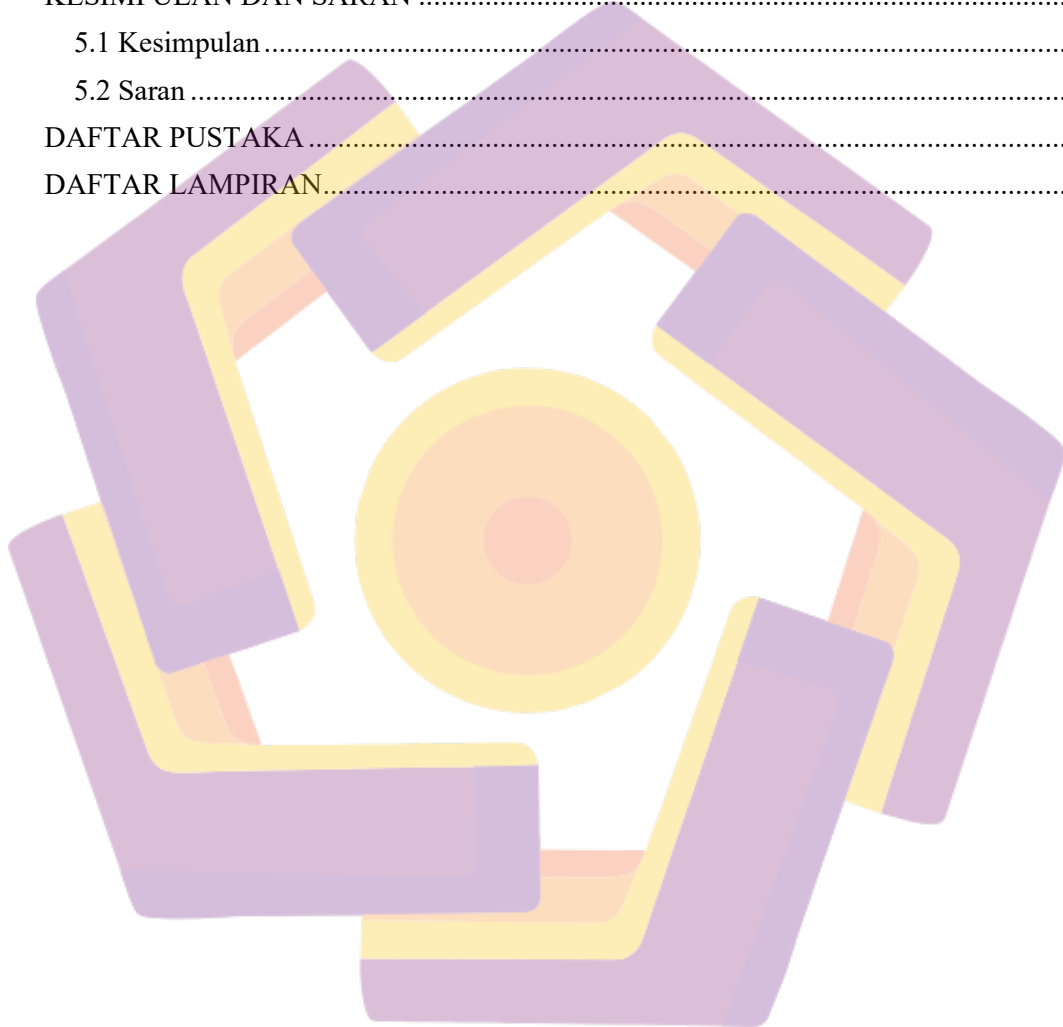
Penulis

DAFTAR ISI

HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
KATA PENGANTAR	v
DAFTAR ISI	vi
DAFTAR TABEL.....	ix
DAFTAR GAMBAR.....	x
DAFTAR LAMPIRAN.....	xii
DAFTAR SIMBOL	xiii
INTISARI	xv
ABSTRACT	xvi
BAB I.....	1
PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	3
1.3 Batasan Masalah.....	3
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	4
1.5.1 Manfaat Teoritis.....	4
1.5.2 Manfaat praktis	5
1.6 Sistematika Penulisan.....	5
BAB II	7
TINJAUAN PUSTAKA	7
2.1 Studi Literatur.....	7
2.2 Dasar Teori	17
2.2.1 Analisis Sentimen	17
2.2.2 Natural Language Processing.....	17
2.2.3 Text Mining.....	18
2.2.4 Data Mining	18
2.2.5 Google Maps Review	19
2.2.6 Web Scraping.....	20
2.2.7 Imbalanced Data	21
2.2.8 Pelabelan Data.....	23
2.2.9 Algoritma Naive Bayes.....	24

2.2.10	Term Frequency-Inverse Document Frequency (TF-IDF)	26
2.2.11	Evaluasi Model	27
2.2.12	Streamlit	31
2.2.13	Skala Likert	32
BAB III		34
METODOLOGI PENELITIAN		34
3.1	Objek Penelitian	34
3.2	Alur Penelitian	35
3.2.1	Scraping Data	35
3.2.2	Cleaning awal data	36
3.2.3	Preprocessing data	37
3.2.4	Pelabelan Data	38
3.2.5	Visualisasi Data : WordCloud dan Review Paling Representatif	39
3.2.6	Ekstraksi fitur (TF-IDF)	39
3.2.7	Split data	41
3.2.8	Handling Imbalancing Data dan training	41
3.2.9	Hyperparameter Tuning	42
3.2.10	Evaluasi Model	42
3.2.11	Test Real Case Model	43
3.2.12	Simpan Model	43
3.2.13	Implementasi Sistem	43
3.3	Alat Dan Bahan	44
BAB IV		46
HASIL DAN PEMBAHASAN		46
4.1	Pengumpulan Data	46
4.1.1	Scraping Pertama	46
4.1.2	Scraping kedua	48
4.2	Pengolahan Data	49
4.2.1	Cleaning Awal	49
4.2.2	Preprocessing Data	51
4.2.3	Pelabelan Data	54
4.2.4	Wordcloud Dan Review Paling Representatif	55
4.2.5	Ekstraksi Fitur (TF-IDF)	60
4.2.6	Split Data	61
4.2.7	Handling Imbalance (SMOTE & ADASYN)	62
4.2.8	Evaluasi Model	67

4.2.9 Test Case Nyata	69
4.2.10 Simpan Model	71
4.3 Implikasi dan Rekomendasi layanan	72
4.4 Implementasi Sistem	73
4.5 Uji Penerimaan Pengguna (<i>User Acceptance Test</i>)	74
BAB V	78
KESIMPULAN DAN SARAN	78
5.1 Kesimpulan	78
5.2 Saran	79
DAFTAR PUSTAKA	80
DAFTAR LAMPIRAN	84



DAFTAR TABEL

Tabel 2. 1 Keaslian Penelitian.....	14
Tabel 3. 1 Perangkat Penellitian.....	45
Tabel 3. 2 Bahan Penelitian.....	45
Tabel 4. 1 Contoh Hasil Case Folding	52
Tabel 4. 2 Contoh Hasil Normalisasi	52
Tabel 4. 3 Contoh Hasil Cleaning	52
Tabel 4. 4 Contoh Hasil Handling Negation	53
Tabel 4. 5 Contoh Hasil Stopword Removal.....	53
Tabel 4. 6 Contoh Hasil Stemming	54
Tabel 4. 7 Pembagian Pelabelan data.....	55
Tabel 4. 8 Perbandingan metode	64
Tabel 4. 9 Perbandingan SMOTE dengan SMOTE+Tuning	68
Tabel 4. 10 Profil Responden UAT	75
Tabel 4. 11 Hasil UAT Berdasarkan Aspek Penilaian	77

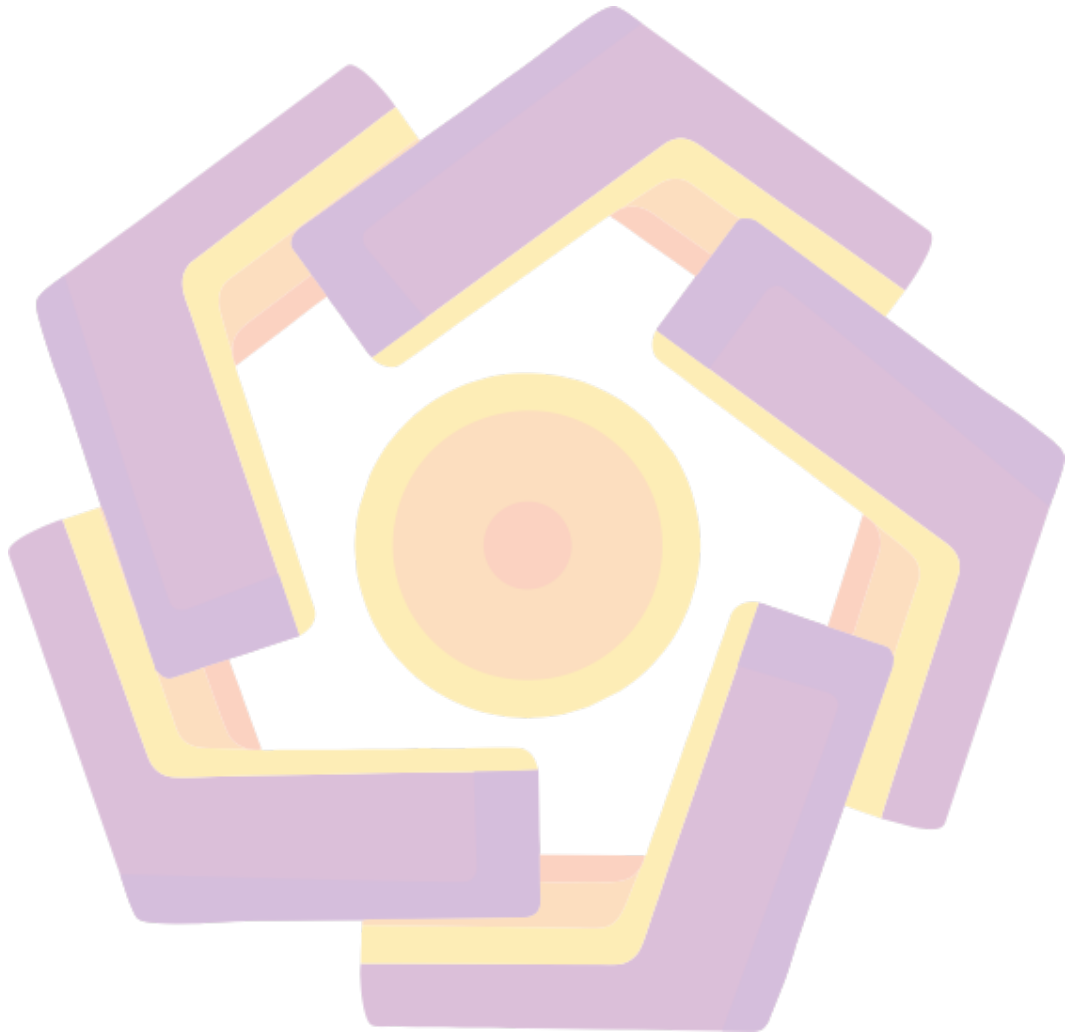
DAFTAR GAMBAR

Gambar 3. 1 Perusahaan Super Dazzle	34
Gambar 3. 2 Alur Penelitian.....	35
Gambar 4. 1 Hasil Akhir Scraping Pertama dan Kedua	49
Gambar 4. 2 Hasil Cleaning Awal.....	51
Gambar 4. 3 Hasil Proses Pelabelan Data	55
Gambar 4. 4 Hasil WordCloud.....	56
Gambar 4. 5 Hasil Review Paling Representatif.....	58
Gambar 4. 6 Hasil proses TF_IDF	61
Gambar 4. 7 Hasil pembagian data (split data)	62
Gambar 4. 8 Hasil Proses Handling Imbalance Data	63
Gambar 4. 9 Hasil Perbandingan Evaluasi Model	64
Gambar 4. 10 Hasil Tunning Pada Model ADASYN	66
Gambar 4. 11 Hasil Tunning Pada Model SMOTE	66
Gambar 4. 12 Hasil Test Case	70
Gambar 4. 13 Hasil Simpan Model.....	72
Gambar 4. 14 Tampilan Aplikasi menampilkan sentimen positif.....	74
Gambar 4. 15 Proses Komunikasi Dengan Pihak Perusahaan	76
Gambar 4. 16 Bukti Pengisian Form UAT.....	76
Gambar L. 1 Skrip File Config.py	85
Gambar L. 2 Skrip File main.py.....	85
Gambar L. 3 Fungsi Utama Pada Skrip scraper.py	86
Gambar L. 4 fungsi get_rating() pada skrip scraper.py	87
Gambar L. 5 fungsi get_full_text() pada skrip scraper.py	88
Gambar L. 6 Proses scraping pertama.....	89
Gambar L. 7 fungsi utama yang diubah pada skrip scrape.py	90
Gambar L. 8 Proses scraping kedua	91
Gambar L. 9 Skrip Cleaning Awal.....	92
Gambar L. 10 Skrip Proses Preprocessing	93
Gambar L. 11 Skrip Case Folding.....	94

Gambar L. 12 Skrip Normalisasi	94
Gambar L. 13 Skrip Cleaning	94
Gambar L. 14 Skrip Handling Negation	95
Gambar L. 15 Skrip Stopword Removal.....	96
Gambar L. 16 Skrip Stemming	97
Gambar L. 17 Hasil Proses Preprocessing	98
Gambar L. 18 Skrip Proses Pelabelan Data	99
Gambar L. 19 Skrip WordCloud.....	100
Gambar L. 20 Skrip Review Paling Representatif.....	101
Gambar L. 21 Skrip Proses TF-IDF	102
Gambar L. 22 Skrip Proses Split Data	102
Gambar L. 23 Skrip Handling Imbalance data.....	103
Gambar L. 24 Skrip Proses Perbandingan	104
Gambar L. 25 Skrip Tuning Pada Model ADASYN	105
Gambar L. 26 Skrip Pada Tuning Model SMOTE	106
Gambar L. 27 Skrip Evaluasi Model.....	107
Gambar L. 28 Hasil Evaluasi Model.....	108
Gambar L. 29 Skrip Test Case	109
Gambar L. 30 Skrip Penyimpanan Model.....	109

DAFTAR LAMPIRAN

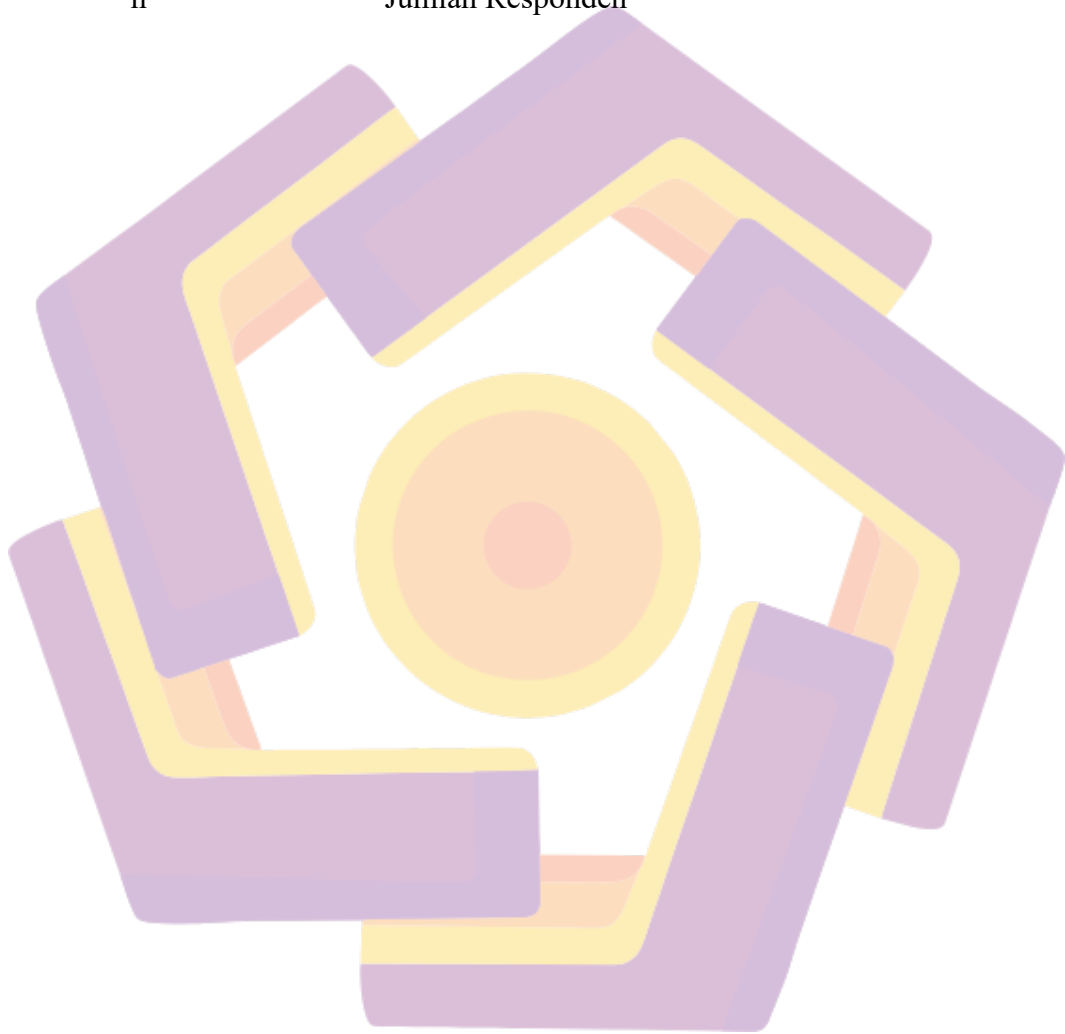
Lampiran 1 Surat izin Penelitian	84
Lampiran 2 Gambar Proses Scraping	85
Lampiran 3 Gambar Proses Pengolahan Data	92



DAFTAR SIMBOL

x_{baru}	data sintetis yang dihasilkan
x_i	data minoritas awa
x_{zi}	tetangga terdekat dari data minoritas
$rand(0,1)$	bilangan acak antara 0 dan 1
G_i	Jumlah data sintetis yang dibuat untuk data minoritas i
r_i	tingkat kesulitan klasifikasi pada data minoritas i
$\sum r_i$	Total tingkat kesulitan seluruh data minoritas
G	Jumlah total data sintetis yang akan dihasilkan
$P(C X)$	Probabilitas suatu kelas C terhadap data X
$P(X C)$	Probabilitas data X pada kelas C
$P(C)$	Probabilitas awal suatu kelas
$P(X)$	Probabilitas data yang diamati
$P(d c)$	Probabilitas dokumen terhadap kelas
$P(c)$	Probabilitas awal kelas
$P(t_k c)$	Probabilitas kemunculan kata pada kelas
$f_{k,d}$	Frekuensi kemunculan kata dalam dokumen
tf	<i>Term Frequency</i> atau banyak kata pada dokumen
$tf_{(t,d)}$	<i>Term Frequency</i> banyaknya kata t pada dokumen d atau pembobotan lokal
$idf_{(t)}$	<i>inverse document frequency</i>
N	Jumlah keseluruhan dokumen
$df_{(t)}$	Jumlah dokumen yang memuat <i>term</i> t
$W_{(t,d)}$	Bobot <i>TF-IDF</i> dari term t pada dokumen d
$tf_{(t,d)}$	Jumlah t dalam d / jumlah kata pada d
$df_{(t)}$	jumlah dokumen yang memuat term
TP	True Positive

TN	True Negative
FP	False Positive
FN	False Negative
$\bar{x}$	Nilai Rata-Rata
$\sum x$	Jumlah Seluruh Skor Responden
n	Jumlah Responden



INTISARI

Ulasan pelanggan menjadi sumber informasi bagi Perusahaan untuk mengetahui tingkat kepuasan pelanggan terhadap kualitas pelayanan yang diberikan. Namun, banyaknya jumlah ulasan membuat proses analisis manual menjadi kurang efektif dan membutuhkan waktu yang cukup lama. Selain itu, isi ulasan yang beragam menyebabkan Perusahaan kesulitan dalam memahami kecenderungan sentimen pelanggan secara cepat dan tepat. Permasalahan tersebut juga ditemukan pada toko retail Super Dazzle yang memiliki banyak ulasan pelanggan dengan berbagai opini terkait pelayanan dan kualitas toko. Penelitian ini bertujuan untuk melakukan analisis sentimen ulasan pelanggan Super Dazzle menggunakan algoritma *Multinomial Naïve Bayes* untuk peningkatan kualitas layanan. Data penelitian ini diperoleh melalui proses *web scraping Google Maps* menggunakan *Playwright*. Tahapan penelitian meliputi *preprocessing data*, *handling negation*, pelabelan data berdasarkan rating, ekstraksi fitur menggunakan *TF-IDF*, penanganan *imbalanced data* menggunakan *SMOTE* dan *ADASYN*, serta *hyperparameter Tuning* untuk meningkatkan performa model. Evaluasi model dilakukan menggunakan *accuracy*, *precision*, *recall*, dan *F1-Score*. Penelitian ini berhasil menunjukkan bahwa algoritma *Multinomial Naïve Bayes* mampu mengklasifikasi sentiment ulasan pelanggan dengan baik. Berdasarkan Hasil pengujian, metode *SMOTE* menghasilkan performa terbaik dibandingkan *ADASYN* dengan Tingkat akurasi 81%. Hyperparameter tuning juga dilakukan pada metode *SMOTE* yang dipilih dan berhasil mendapatkan akurasi 94,8% pada *alpha* 0.01. Hasil akhir Evaluasi dari metode *SMOTE* yang telah di *tuning* mendapatkan *accuracy* 81%, *precision* 81%, *recall* 81%, dan *F1-Score* 81%. Model yang telah dibangun kemudian diimplementasikan ke dalam aplikasi berbasis web menggunakan *ssstreamlit* untuk mempermudah proses klasifikasi sentimen secara otomatis.

Kata kunci: analisis sentimen, *Naïve Bayes*, *TF-IDF*, *SMOTE*, Google Maps.

ABSTRACT

Customer reviews are a source of information for companies to determine the level of customer satisfaction with the quality of service provided. However, the large number of reviews makes the manual analysis process less effective and requires a considerable amount of time. In addition, the diverse content of reviews makes it difficult for companies to quickly and accurately understand customer sentiment trends. This problem was also found in the Super Dazzle retail store, which has many customer reviews with various opinions regarding service and store quality. This study aims to conduct a sentiment analysis of Super Dazzle customer reviews using the Multinomial Naïve Bayes algorithm to improve service quality. This research data was obtained through the Google Maps web scraping process using Playwright. The research stages include data preprocessing, handling negation, data labeling based on ratings, feature extraction using TF-IDF, handling imbalanced data using SMOTE and ADASYN, and hyperparameter tuning to improve model performance. Model evaluation was carried out using accuracy, precision, recall, and F1-Score. This study successfully demonstrated that the Multinomial Naïve Bayes algorithm is able to classify customer review sentiment well. Based on the test results, the SMOTE method produced the best performance compared to ADASYN with an accuracy level of 81%. Hyperparameter tuning was also performed on the selected SMOTE method, achieving 94.8% accuracy at an alpha of 0.01. The final evaluation results of the tuned SMOTE method yielded 81% accuracy, 81% precision, 81% recall, and an F1-score of 81%. The developed model was then implemented into a web-based application using sstreamlit to simplify the automatic sentiment classification process.

Keywords: sentiment analysis, Naïve Bayes, TF-IDF, SMOTE, Google Maps.

-