

**ANALISIS SENTIMEN TERHADAP REVIEW PENGGUNAAN
PRODUK SUNSCREEN DENGAN MENGGUNAKAN
ALGORITMA NAIVE BAYES CLASSIFIER**

SKRIPSI

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



disusun oleh

MUTIA ZAHIRMA

21.11.4205

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

**ANALISIS SENTIMEN TERHADAP REVIEW PENGGUNAAN
PRODUK SUNSCREEN DENGAN MENGGUNAKAN
ALGORITMA NAIVE BAYES CLASSIFIER**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh
MUTIA ZAHIRMA
21.11.4205

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

HALAMAN PERSETUJUAN

SKRIPSI

**ANALISIS SENTIMEN TERHADAP REVIEW PENGGUNAAN PRODUK
SUNSCREEN DENGAN MENGGUNAKAN ALGORITMA NAIVE BAYES
CLASSIFIER**

yang disusun dan diajukan oleh

Mutia Zahirma

21.11.4205

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 23 Januari 2025

Dosen Pembimbing,



Rumini, M. Kom
NIK. 190302246

HALAMAN PENGESAHAN

SKRIPSI

ANALISIS SENTIMEN TERHADAP REVIEW PENGGUNAAN PRODUK
SUNSCREEN DENGAN MENGGUNAKAN ALGORITMA NAIVE BAYES
CLASSIFIER

yang disusun dan diajukan oleh

Mutia Zahirma

21.11.4205

Telah dipertahankan di depan Dewan Penguji
pada tanggal 23 Januari 2025

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Nur Aini, A.Md., S.Kom., M.Kom.
NIK. 190302066



Supriatin, A.Md., S.Kom., M.Kom.
NIK. 190302239



Rumini, S.Kom., M.Kom.
NIK. 190302246



Skrripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 23 Januari 2025

DEKAN FAKULTAS ILMU KOMPUTER



Hanif Al Fatta, S.Kom., M.Kom., Ph.D.
NIK. 190302096

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Mutia Zahirma

NIM : 21.11.4205

Menyatakan bahwa Skripsi dengan judul berikut:

ANALISIS SENTIMEN TERHADAP REVIEW PENGGUNAAN PRODUK SUNSCREEN DENGAN MENGGUNAKAN ALGORITMA NAIVE BAYES CLASSIFIER

Dosen Pembimbing : Rumini, M. Kom

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, 23 Januari 2025

Yang Menyatakan,



Mutia Zahirma

HALAMAN PERSEMBAHAN

Puji syukur yang selalu penulis panjatkan ke hadirat Allah SWT, karena atas karunia dan limpahan rahmat yang diberikan sehingga penulis dapat menyelesaikan skripsi yang berjudul “Analisis Sentimen Terhadap Review Penggunaan Produk Sunscreen Dengan Menggunakan Algoritma Naive Bayes Classifier”. Dalam perjalanan panjang ini, penulis merasa beruntung mendapatkan dukungan dari berbagai pihak. Oleh karena itu, dengan tulus penulis mempersembahkan skripsi ini kepada :

1. Diri saya sendiri, yang telah berjuang untuk menyelesaikan skripsi ini meskipun banyak rintangan yang dihadapi dan rasa tidak percaya diri yang selalu ada.
2. Orang tua penulis, yang selalu menjadi tempat pertama untuk penulis berkeluh kesah. Terimakasih telah memberikan doa dan dukungan kepada penulis yang tiada henti.
3. Saudari penulis, Karina Septiani Marzuni, Zaskia Zaharani, dan Amanda Rafadina yang selalu mendukung dan menghibur penulis saat merasa sedih.
4. Ibu Rumini, M.Kom, selaku dosen pembimbing penulis yang telah membimbing dan memberikan arahan selama penyusunan skripsi ini.
5. Fitrah Islami Ranasti, yang selalu menemani dan memberikan dukungan kepada penulis.
6. Teman-teman IF 06 dan para sahabat, yang sudah bersama-sama berjuang dari awal masa perkuliahan. Terimakasih sudah saling memberikan dukungan dan canda tawa.
7. Kampus Ungu tercinta, tempat di mana penulis mendapatkan pendidikan yang berkualitas dan pengalaman yang tak ternilai. Terima kasih kepada seluruh civitas akademika yang telah mendukung proses pembelajaran ini.

KATA PENGANTAR

Puji syukur penulis panjatkan kepada Tuhan Yang Maha Esa, karena berkat rahmat dan karunia-Nya, penulis dapat menyelesaikan skripsi ini dengan judul “Analisis Sentimen Terhadap Review Penggunaan Produk Sunscreen Dengan Menggunakan Algoritma Naive Bayes Classifier”. Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana di Universitas Amikom Yogyakarta.

Penyusunan skripsi ini tidak terlepas dari bantuan, dukungan, dan dorongan dari berbagai pihak. Oleh karena itu, penulis ingin menyampaikan rasa terima kasih yang sebesar-besarnya kepada:

1. Bapak Prof. Dr. M. Suyanto, MM selaku Rektor Universitas Amikom Yogyakarta.
2. Bapak Hanif Al Fatta, S.Kom., M.Kom, Selaku Dekan Fakultas Ilmu Komputer.
3. Ibu Windha Mega Pradnya D, M.Kom selaku ketua program studi SI Informatika.
4. Ibu Rumi, M.Kom, selaku dosen pembimbing saya yang telah banyak memberikan pengarahan bagi penulis dalam pembuatan skripsi.
5. Bapak-Ibu Resource Centre Universitas Amikom Yogyakarta, atas dukungan yang diberikan selama penulis menyusun naskah skripsi.

Penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan. Oleh karena itu, penulis mengharapkan kritik dan saran yang membangun dari pembaca demi perbaikan di masa yang akan datang. Semoga skripsi ini dapat memberikan manfaat dan sumbangsih bagi ilmu pengetahuan, khususnya dalam bidang ilmu pengetahuan dan teknologi.

Yogyakarta, 23 Januari 2025

Mutia Zahirma

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN	ii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iii
HALAMAN PERSEMBAHAN	iv
KATA PENGANTAR	vi
DAFTAR ISI	vii
DAFTAR TABEL	x
DAFTAR GAMBAR	xi
ABSTRACT	xiv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	4
1.6 Sistematika Penulisan	4
BAB II TINJAUAN PUSTAKA	6
2.1 Studi Literatur	6
2.2 Dasar Teori	13
2.2.1 <i>Machine Learning</i>	13
2.2.2 <i>Data Mining</i>	13

2.2.3	Analisis Sentimen	13
2.2.4	<i>Naive Bayes Classifier</i>	14
2.2.5	<i>Confusion Matrix</i>	15
2.2.6	Pembagian Dataset.....	16
2.2.7	<i>Sunscreen Wardah UV Shield Essential Gel SPF 35 PA +++</i>	17
2.2.8	<i>Femaledaily</i>	17
2.2.9	<i>Google Collaboratory</i>	17
2.2.10	<i>GitHub</i>	18
2.2.11	<i>Streamlit</i>	18
BAB III METODE PENELITIAN		19
3.1	Objek Penelitian.....	19
3.2	Alur Penelitian	19
3.2.1	Pengambilan Data	21
3.2.2	Data Labeling.....	21
3.2.3	<i>Preprocessing Data</i>	22
3.2.4	<i>Transformation(TF-IDF)</i>	23
3.2.5	Pembagian Data	24
3.2.6	<i>Data Mining (Naive Bayes Classifier)</i>	24
3.2.7	<i>Evaluation Confusion Matrix</i>	25
3.2.8	<i>Deployment Aplikasi</i>	26
3.3	Alat dan Bahan	28
3.3.1	Data Penelitian	28
3.3.2	Alat/instrumen.....	28
BAB IV HASIL DAN PEMBAHASAN		29
4.1	Pengumpulan Data	29

4.2	Pelabelan Data	31
4.3	<i>Preprocessing Data</i>	32
4.3.1	<i>Cleaning Dataset</i>	32
4.3.2	<i>Case Folding</i>	33
4.3.3	<i>Tokenizing</i>	36
4.3.4	<i>Stopword Removal</i>	37
4.4	<i>Transformation TF-IDF</i>	38
4.5	Pembagian Data	39
4.6	<i>Data Mining (Naive Bayes Classifier)</i>	39
4.7	<i>Evaluation Confusion Matrix</i>	40
4.8	<i>Deployment Aplikasi</i>	42
BAB V PENUTUP		46
5.1	Kesimpulan	46
5.2	Saran	47
REFERENSI		48

DAFTAR TABEL

Tabel 2. 1 Keaslian Penelitian	8
Tabel 3. 1 Data Labeling.....	21
Tabel 3. 2 Case Folding.....	22
Tabel 3. 3 Tokenizing	23
Tabel 3. 4 Stopword Removal	23
Tabel 3. 5 TF-IDF	24
Tabel 3. 6 Pembagian Data	24
Tabel 3. 7 Klasifikasi Data (Pembuatan Model).....	25
Tabel 3. 8 Representasi dari Confusion Matrix 2x2	26
Tabel 3. 9 Deployment Aplikasi	27

DAFTAR GAMBAR

Gambar 2. 1 Struktur Confusion Matrix (untuk klasifikasi biner).....	15
Gambar 3. 1 Alur Penelitian	20
Gambar 3. 2 Wardah UV Shield Essential Sunscreen Gel SPF 35 PA +++.....	21
Gambar 4. 1 Install Selenium & BeautifulSoup	29
Gambar 4. 2 Proses Scraping Data	30
Gambar 4. 3 Proses Scraping Data	30
Gambar 4. 4 Proses Pengumpulan Data.....	31
Gambar 4. 5 Hasil Scraping Data	31
Gambar 4. 6 Proses Pelabelan Data	31
Gambar 4. 7 Hasil Pelabelan Data	32
Gambar 4. 8 Jumlah Sentimen Positif dan Negatif.....	32
Gambar 4. 9 Menghapus Missing Values	32
Gambar 4. 10 Hasil Penghapusan Missing Values	33
Gambar 4. 11 Menghapus Duplikat Data	33
Gambar 4. 12 Hasil Penghapusan Duplikat Data.....	33
Gambar 4. 13 Proses Case Folding	34
Gambar 4. 14 Proses Remove Digits	34
Gambar 4. 15 Proses Remove Punctuation.....	35
Gambar 4. 16 Hasil Proses Case Folding.....	35
Gambar 4. 17 Hasil Proses Remove Digits.....	35
Gambar 4. 18 Hasil Proses Remove Punctuation	35
Gambar 4. 19 Proses Tokenizing	36
Gambar 4. 20 Hasil Tokenizing	37
Gambar 4. 21 Proses Stopword Removal	37
Gambar 4. 22 Hasil Stopword Removal	37
Gambar 4. 23 Proses TF-IDF	38

Gambar 4. 24 Proses TF-IDF	38
Gambar 4. 25 Proses Pembagian Data	39
Gambar 4. 26 Proses Pembagian Data	39
Gambar 4. 27 Proses Naive Bayes Classifier.....	39
Gambar 4. 28 Proses Klasifikasi Report	40
Gambar 4. 29 Hasil Kalsifikasi Report	40
Gambar 4. 30 Hasil Evaluation Confusion Matrix	41
Gambar 4. 31 Visualisasi WordCloud Sentimen Positif.....	42
Gambar 4. 32 Visualisasi WordCloud Sentimen Negatif	42
Gambar 4. 33 Proses Menyimpan Model	43
Gambar 4. 34 Kode Program Aplikasi Streamlit.....	43
Gambar 4. 35 Kode Program Aplikasi Streamlit	44
Gambar 4. 36 Kode Program Aplikasi Streamlit	44
Gambar 4. 37 Hasil Ulasan Positif.....	45
Gambar 4. 38 Hasil Ulasan Negatif	45

INTISARI

Kemajuan teknologi informasi dan komunikasi telah mengubah interaksi konsumen dengan produk dan merek, terutama di industri kecantikan. Penelitian ini berfokus pada analisis sentimen ulasan produk sunscreen dengan metode Naive Bayes Classifier. Data ulasan produk *Wardah UV Shield Essential Sunscreen Gel SPF 35 PA+++* dikumpulkan melalui *web scraping* dari situs *Femaledaily*, menghasilkan 1450 baris data. Data tersebut diberi label sentimen positif atau negatif berdasarkan rating, kemudian diproses melalui tahap pembersihan data, case folding, tokenisasi, dan penghapusan kata umum (stopword removal). Data yang telah diproses kemudian diubah menjadi representasi numerik menggunakan *TF-IDF*. Model *Naive Bayes Classifier* yang dibangun berhasil mencapai akurasi 84%, precision 83%, recall 83%, dan F1-score 83%. Visualisasi dengan word cloud menunjukkan kata-kata yang sering muncul dalam ulasan positif dan negatif. Penelitian ini menunjukkan bahwa metode *Naive Bayes Classifier* efektif untuk mengklasifikasikan sentimen ulasan produk sunscreen. Meskipun metode ini mudah diimplementasikan dan dipahami, terdapat kelemahan dalam asumsi independensi kata dan ketidakseimbangan jumlah ulasan positif dan negatif. Penelitian selanjutnya diharapkan dapat memperluas cakupan dataset dan mempertimbangkan metode analisis sentimen lain untuk meningkatkan akurasi.

Kata kunci: Sunscreen, Wardah, Analisis Sentimen, Ulasan Konsumen, Naive Bayes Classifier.

ABSTRACT

Advancements in information and communication technology have transformed consumer interactions with products and brands, particularly in the beauty industry. This study focuses on sentiment analysis of sunscreen product reviews using the Naive Bayes Classifier method. Reviews for the Wardah UV Shield Essential Sunscreen Gel SPF 35 PA+++ were collected through web scraping from the Femaledaily website, resulting in 1450 data entries. These reviews were labeled as either positive or negative based on their ratings, then underwent data preprocessing, including data cleaning, case folding, tokenization, and stopword removal. The processed data was then converted into numerical representations using TF-IDF. The Naive Bayes Classifier model developed in this study achieved an accuracy of 84%, a precision of 83%, a recall of 83%, and an F1-score of 83%. Visualization with word clouds highlighted frequently occurring words in positive and negative reviews. This research demonstrates that the Naive Bayes Classifier method is effective for classifying sentiment in sunscreen product reviews. Despite its ease of implementation and understanding, the method has limitations, such as the assumption of word independence and the imbalance between the number of positive and negative reviews. Future research is expected to expand the dataset's scope and consider other sentiment analysis methods to improve accuracy.

Keyword: Sunscreen, Wardah, Sentiment Analysis, Consumer Reviews, Naive Bayes Classifier.