



**ANALISIS SENTIMEN PADA FORUM *REVIEW* FEMALE DAILY
TERHADAP PRODUK SKINTIFIC MENGGUNAKAN METODE
KLASIFIKASI NAÏVE BAYES**

SKRIPSI

AQILA ZAHARA PUTRI

NIM. 2010512071

S1 SISTEM INFORMASI

FAKULTAS ILMU KOMPUTER

UNIVERSITAS PEMBANGUNAN NASIONAL “VETERAN” JAKARTA

2024



**ANALISIS SENTIMEN PADA FORUM *REVIEW* FEMALE DAILY
TERHADAP PRODUK SKINTIFIC MENGGUNAKAN METODE
KLASIFIKASI NAÏVE BAYES**

SKRIPSI

**Diajukan Sebagai Salah Satu Syarat untuk Memperoleh Gelar
Sarjana Komputer**

AQILA ZAHARA PUTRI

NIM. 2010512071

S1 SISTEM INFORMASI

FAKULTAS ILMU KOMPUTER

UNIVERSITAS PEMBANGUNAN NASIONAL “VETERAN” JAKARTA

2024

PERNYATAAN ORISINALITAS

Skripsi ini adalah hasil karya sendiri, dan semua sumber yang dikutip maupun dirujuk telah saya nyatakan dengan benar.

Nama : Aqila Zahara Putri

NIM : 2010512071

Tanggal : 5 Juli 2024

Bilamana di kemudian hari ditemukan ketidaksesuaian dengan pernyataan ini, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan berlaku.

Jakarta, 5 Juli 2024

Yang Menyatakan,



DF20ALX179445622
(Aqila Zahara Putri)

**PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI
UNTUK KEPENTINGAN AKADEMIS**

Sebagai civitas akademik Universitas Pembangunan Nasional “Veteran” Jakarta, saya yang bertanda tangan di bawah ini:

Nama : Aqila Zahara Putri

NIM : 2010512071

Fakultas : Ilmu Komputer

Program Studi : S1 – Sistem Informasi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Pembangunan Nasional “Veteran” Jakarta Hak Bebas Royalti Non-Eksklusif (*Non-Exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul:

Analisis Sentimen Pada Forum Review Female Daily Terhadap Produk Skintific Menggunakan Metode Klasifikasi Naïve Bayes

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti ini Universitas Pembangunan Nasional “Veteran” Jakarta berhak menyimpan, mengalih data/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat dan mempublikasikan Tugas Akhir saya selama tetap mencantumkan nama saya sebagai pemilik Hak Cipta. Demikian pernyataan ini saya buat sebenarnya.

Dibuat di : Jakarta
Pada Tanggal : 5 Juli 2024
Yang Menyatakan,



(Aqila Zahara Putri)

LEMBAR PENGESAHAN

Skripsi ini diajukan oleh :

Nama : Aqila Zahara Putri

NIM : 2010512071

Program Studi : S1 – Sistem Informasi

Judul Tugas Akhir : Analisis Sentimen Pada Forum *Review* Female Daily Terhadap Produk Skintific Menggunakan Metode Klasifikasi Naïve Bayes

Telah berhasil dipertahankan dihadapan Tim Penguji dan diterima sebagai bagian dari persyaratan yang diperlukan untuk memperoleh gelar Sarjana Ilmu Komputer pada Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional “Veteran” Jakarta.



Ridwan Raafi'udin, S.Kom., M.Kom.

Penguji 1



Nindy Irzavika, S.SI., M.TI.

Penguji 2



Nurhafifah Matondang, S.Kom., M.M., M.TI

Pembimbing 1



Anita Muliawati, S.Kom., M.TI.

Pembimbing 2



Prof. Dr. Ir. Supriyanto, ST., M.Sc., IPM

Dekan



Anita Muliawati, S.Kom., M.TI.

Ketua Program Studi S1 Sistem Informasi

Ditetapkan di : Jakarta

Tanggal Ujian : 10 Juni 2024

ANALISIS SENTIMEN PADA FORUM *REVIEW* FEMALE DAILY TERHADAP PRODUK SKINTIFIC MENGGUNAKAN METODE KLASIFIKASI NAIVE BAYES

Aqila Zahara Putri

ABSTRAK

Kini media sosial telah menjadi sarana yang paling diandalkan dalam memperoleh suatu informasi dan berinteraksi dengan sifat komunikasi dua arah. Dalam konteks industri kecantikan, media sosial menjadi platform yang sangat berpengaruh dalam menyebarkan informasi mengenai produk-produknya. Peningkatan pesat popularitas *brand* Skintific merupakan salah satu dampak positif media sosial terhadap industri kecantikan. Skintific 5x Ceramide Barrier Repair Moisture Gel Moisturizer merupakan produknya yang menjadi sorotan masyarakat karena memiliki formulasi yang berfokus dalam meningkatkan fungsi *skin barrier*. Meskipun demikian, terdapat pengguna yang melaporkan tidak cocok bahkan mengalami reaksi kulit yang tidak diinginkan. Oleh karena itu, sebelum memutuskan menggunakan suatu produk, masyarakat melakukan berbagai riset, seperti membaca *review* produk. Namun, banyaknya *review* membuat masyarakat membutuhkan waktu lebih untuk memutuskan membeli suatu produk. Sehingga pada penelitian ini dilakukan analisis sentimen pengguna forum *review* Female Daily terhadap produk Skintific 5x Ceramide Barrier Repair Moisture Gel Moisturizer, menggunakan metode klasifikasi Multinomial Naive Bayes. Hasil evaluasi Confusion Matrix menunjukkan model memiliki tingkat akurasi yang baik, dengan nilai *accuracy* sebesar 87,74%, *precision* sebesar 96,25%, *recall* sebesar 89,86%, dan *F1 score* sebesar 0,9295. Hasil penelitian ini juga dianalisis dengan pembuatan visualisasi pada Tableau Public.

Kata Kunci: Analisis Sentimen, Female Daily, Skintific, Naive Bayes, TF-IDF, Tableau

**SENTIMENT ANALYSIS ON FEMALE DAILY REVIEW
FORUM TOWARDS SKINTIFIC PRODUCTS USING
NAIVE BAYES CLASSIFICATION METHOD**

Aqila Zahara Putri

ABSTRACT

Nowadays, social media has become the most relied-upon means of obtaining information and engaging in two-way communication. In the context of the beauty industry, social media serves as a highly influential platform for disseminating information about its products. The rapid increase in the popularity of the Skintific brand is one of the positive impacts of social media on the beauty industry. Skintific's 5x Ceramide Barrier Repair Moisture Gel Moisturizer has garnered public attention due to its formulation focused on enhancing skin barrier function. However, some users have reported dissatisfaction or even adverse skin reactions. Therefore, before deciding to use a product, consumers conduct various research, such as reading product reviews. However, the abundance of reviews makes it time-consuming for consumers to decide on a purchase. Hence, this research conducts sentiment analysis of Female Daily forum users towards Skintific 5x Ceramide Barrier Repair Moisture Gel Moisturizer, using the Multinomial Naive Bayes classification method. Confusion Matrix evaluation results show the model has good accuracy, with an accuracy value of 87.74%, precision of 96.25%, recall of 89.86%, and F1 score of 0.9295. The research findings are also analyzed through visualization on Tableau Public.

Keywords: *Sentiment Analysis, Female Daily, Skintific, Naive Bayes, TF-IDF, Tableau*

KATA PENGANTAR

Puji dan syukur dipanjatkan atas kepada Allah SWT. yang telah memberikan Rahmat serta karunia-Nya, sehingga penulis dapat menyelesaikan Tugas Akhir dengan judul “Analisis Sentimen Pada Forum *Review* Female Daily Terhadap Produk Skintific Menggunakan Metode Klasifikasi Naive Bayes” dengan baik. Selama proses penyelesaian proposal ini, penulis menyadari banyak pihak yang membantu dan berkontribusi baik berupa dukungan moril maupun materiel. Untuk itu penulis mengucapkan banyak terima kasih kepada:

1. Allah SWT yang telah memberikan petunjuk, kesabaran, dan keteguhan kepada penulis sehingga dapat menyelesaikan Tugas Akhir ini dengan baik.
2. Kedua orang tua yang banyak memberikan dukungan baik secara moral dan materiel, dan motivasi selama penyelesaian Tugas Akhir.
3. Bapak Prof. Dr. Ir. Supriyanto, ST., M.Sc., IPM. selaku dekan Fakultas Ilmu Komputer.
4. Ibu Anita Muliawati, S.Kom., M.T.I., selaku Ketua Program Studi Sarjana Jurusan Sistem Informasi serta Dosen Pembimbing 2 yang telah membimbing, memberi saran dan arahan selama penulisan Tugas Akhir.
5. Ibu Nurhafifah Matondang, S.Kom., M.T.I selaku Dosen Pembimbing 1 yang telah membimbing, memberi saran, arahan, dukungan, dan doa selama penulisan Tugas Akhir.
6. Khansa Shahira, Annisa Azzahra, dan Keluarga Besar Alfito selaku sahabat yang selalu mendukung dan membantu sepenuh hati.
7. Pihak-pihak lain yang tidak dapat disebutkan secara satu – persatu.

Penulis menyadari bahwa masih banyak kekurangan pada Tugas Akhir ini. Oleh karena itu, kritik dan saran yang bersifat membangun demi akan sangat berharga bagi penulis.

Jakarta, 15 April 2024

Penulis

DAFTAR ISI

LEMBAR JUDUL	ii
ERNYATAAN ORISINALITAS	iii
PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS.....	iv
LEMBAR PENGESAHAN	v
ABSTRAK	vi
ABSTRACT	vii
KATA PENGANTAR.....	viii
DAFTAR ISI	ix
DAFTAR TABEL.....	xii
DAFTAR GAMBAR	xiv
DAFTAR SIMBOL	xv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	6
1.3 Tujuan Penelitian.....	7
1.4 Manfaat Penelitian.....	7
1.5 Batasan Masalah.....	8
BAB II TINJAUAN PUSTAKA	10
2.1 Produk Skintific.....	10
2.2 Forum <i>Review</i> Female Daily	10
2.3 <i>Data Mining</i>	11
2.3.1 Text Mining.....	12
2.4 Uji Reliabilitas Kappa Fleiss	13
2.5 <i>Pre-processing Data</i>	15
2.5.1 Data Cleaning dan Case Folding.....	15
2.5.2 Normalisasi	16
2.5.3 Stopword Removal.....	16
2.5.4 Stemming	16
2.5.5 Tokenization	17
2.6 Pembobotan TF-IDF (<i>Term Frequency - Inverse Document Frequency</i>)17	

2.7	SMOTE (Synthetic Minor Oversampling Technique)	19
2.8	Metode Klasifikasi Naïve Bayes	19
2.8.1	Multinomial Naïve Bayes	20
2.9	<i>Confusion Matrix</i>	21
2.10	Bahasa Pemrograman Python	22
2.11	Penelitian Terkait	23
2.12	Hipotesis	30
BAB III METODE PENELITIAN.....		31
3.1	Tahapan Penelitian.....	31
3.1.1	Identifikasi Masalah	32
3.1.2	Studi Literatur	32
3.1.3	Pengumpulan Data	32
3.1.4	Pelabelan Data.....	33
3.1.5	Pre-processing Data	34
3.1.6	Pembobotan TF-IDF	37
3.1.7	Splitting data	39
3.1.8	Klasifikasi Multinomial Naïve Bayes	39
3.1.9	Evaluasi	39
3.2	Perangkat Penelitian	39
3.3	Jadwal Penelitian	40
BAB IV HASIL DAN PEMBAHASAN.....		42
4.1	Pengumpulan Data	42
4.2	Pelabelan Data	43
4.3	Pre-processing Data.....	48
4.3.1	Data Cleaning dan Case Folding.....	48
4.3.2	Normalisasi	49
4.3.3	Stopword Removal.....	50
4.3.4	Stemming	51
4.3.5	Tokenization.....	52
4.4	Pembobotan TF-IDF.....	53
4.5	Splitting Data.....	57
4.6	Klasifikasi Naïve Bayes	58
4.7	Evaluasi Hasil Klasifikasi	66
4.8	Visualisasi.....	68

4.8.1	Visualisasi Product Usage.....	70
4.8.2	Visualisasi Trend Review.....	74
4.8.3	Visualisasi Word Cloud.....	76
BAB V KESIMPULAN.....		82
5.1	Kesimpulan.....	82
5.2	Saran.....	83
DAFTAR PUSTAKA.....		85
RIWAYAT HIDUP.....		91
LAMPIRAN.....		92

DAFTAR TABEL

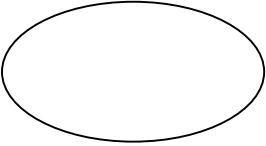

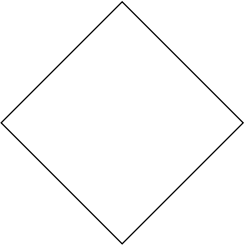
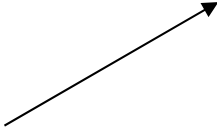
Tabel 2.1 Interpretasi Kappa Fleiss	14
Tabel 2.2 Tabel Confusion Matrix.....	21
Tabel 2.3 Penelitian Terkait.....	23
Tabel 3.1 Contoh Pelabelan Data	33
Tabel 3.2 Contoh Case Folding.....	34
Tabel 3.3 Contoh Normalisasi.....	35
Tabel 3.4 Contoh Stopword Removal	35
Tabel 3.5 Contoh Stemming.....	36
Tabel 3.6 Contoh Tokenization	36
Tabel 3.7 Sampel Data	37
Tabel 3.8 Perhitungan Pembobotan TF-IDF	38
Tabel 3.9 Jadwal Penelitian.....	40
Tabel 4. 1 Deskripsi atribut dari data hasil scraping	43
Tabel 4. 2 Tabel Hasil Pelabelan 3 Anotator	44
Tabel 4. 3 Tabel Hasil Keseluruhan Pelabelan.....	46
Tabel 4. 4 Hasil penerapan Data Cleaning dan Case Folding.....	48
Tabel 4. 5 Hasil Penerapan Normalisasi	50
Tabel 4. 6 Hasil Penerapan Stopword Removal.....	51
Tabel 4. 7 Hasil Penerapan Stemming	52
Tabel 4. 8 Hasil Penerapan Tokenization	53
Tabel 4. 9 Data Sampel	54
Tabel 4. 10 Tabel Perhitungan TF-IDF	54
Tabel 4. 11 Data Setelah Proses Splitting	57
Tabel 4. 12 Data Setelah Proses Splitting dengan SMOTE	58
Tabel 4. 13 Sampel Data Training.....	58
Tabel 4. 14 Nilai Seluruh Likelihood Probability	60
Tabel 4. 15 Sampel Data Testing.....	62
Tabel 4. 16 Sampel Data Testing Setelah Melalui Tahap Pre-processing	62
Tabel 4. 17 Hasil Pembobotan TF-IDF Pada Data Testing.....	63
Tabel 4. 18 Hasil Perhitungan Likelihood Probability.....	64
Tabel 4. 19 Nilai Accuracy tiap Fold	66

Tabel 4. 20 Nilai Evaluasi Tiap Fold.....	66
Tabel 4. 21 Review Positif Dengan Kata "Kulit" dan "Pakai"	77
Tabel 4. 22 Review Positif Dengan Kata "Lembap" dan "Bagus"	78
Tabel 4. 23 Review Negatif Dengan Kata “Jerawat” dan “Beruntus”	79
Tabel 4. 24 Review Negatif Dengan Kata “Produk” dan “Minyak”	80

DAFTAR GAMBAR

Gambar 1. 1 Pangsa Pasar Industri Kecantikan Indonesia.....	1
Gambar 1. 2 Market Share Percentage 10 Best Moisturizer Brand	2
Gambar 1. 3 Review Tidak Cocok Penggunaan Skintific.....	3
Gambar 3. 1 Tahapan Penelitian	31
Gambar 4. 1 Penggunaan Ekstensi Data Miner	42
Gambar 4. 2 Hasil Scraping	42
Gambar 4. 3 Contoh Data Review Tidak Relevan dan Duplikasi.....	46
Gambar 4. 4 Visualisasi Total Jumlah Sentimen.....	48
Gambar 4. 5 Hasil TF-IDF Pada Python.....	57
Gambar 4. 6 Hasil Confusion Matrix.....	67
Gambar 4. 7 Dashboard 1 Visualisasi Sentimen Produk Skintific.....	69
Gambar 4. 8 Dashboard 2 Visualisasi Word Cloud.....	69
Gambar 4. 9 Visualisasi Product Usage Page Less Than 1 Week.....	70
Gambar 4. 10 Visualisasi Product Usage Page 1 Week – 1 Month.....	71
Gambar 4. 11 Visualisasi Product Usage Page 1 Month – 3 Months.....	71
Gambar 4. 12 Visualisasi Product Usage Page 3 Months – 6 Months.....	72
Gambar 4. 13 Visualisasi Product Usage Page 6 Months – 1 Year	73
Gambar 4. 14 Visualisasi Product Usage Page More Than 1 Year	73
Gambar 4. 15 Visualisasi Trend Review Pada Tahun 2021	74
Gambar 4. 16 Visualisasi Trend Review Page 2022	75
Gambar 4. 17 Visualisasi Trend Review Page 2023	76
Gambar 4. 18 Visualisasi Word Cloud Positif.....	77
Gambar 4. 19 Visualisasi Word Cloud Negatif	79

DAFTAR SIMBOL

No.	Simbol Flowchart	Nama	Arti Simbol Flowchart
1		Terminal	Mengindikasikan titik awal atau akhir dari sebuah flowchart.
2		Proses	Merepresentasikan suatu tahapan dalam sebuah proses.
3.		Decision	Mengindikasikan suatu kondisi yang akan menghasilkan dua kemungkinan, yaitu Ya atau Tidak
4.		<i>Flowline</i>	Mengindikasikan arah dari sebuah proses. Tiap <i>flowline</i> akan terhubung pada dua buah blok proses