

LAPORAN TUGAS AKHIR

ANALISIS SENTIMEN PRODUK PEMBERSIH WAJAH
MENGUNAKAN METODE GAUSSIAN NAÏVE BAYES



Disusun Oleh:

Preity Deldiana Putri

NIM: 04221024

PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS NAROTAMA

LEMBAR PENGESAHAN

LAPORAN TUGAS AKHIR

ANALISIS SENTIMEN PRODUK PEMBERSIH
WAJAH MENGGUNAKAN METODE GAUSSIAN
NAÏVE BAYES

DISUSUN OLEH:

PREITY DELDIANA PUTRI

NIM: 04221024

Dipertahankan di depan Penguji Ujian Tugas Akhir

Program Studi Sistem Informasi

Fakultas Ilmu Komputer

Universitas Narotama Surabaya

Tanggal 24 Juli 2025

PRO PATRIA

Penguji

Ketua Program Studi

1. Eman Setiawan, S.Kom., M.M.
NIDN: 0720017501

Eman Setiawan, S.Kom., M.M.
NIDN: 0720017501

2. Dr. Aryo Nugroho, S.T., S.Kom., M.T.
NIDN: 0721077001

Fakultas Ilmu Komputer
Dekan

3. Awalludiyah Ambarwati, S.Kom., M.M.
NIDN: 0015047801

Dr. Carvo Darujati, S.T., M.T.
NIDN: 0710097402

LEMBAR PERSETUJUAN

LAPORAN TUGAS AKHIR

ANALISIS SENTIMEN PRODUK PEMBERSIH WAJAH MENGGUNAKAN METODE GAUSSIAN NAÏVE BAYES

DISUSUN OLEH:

PREITY DELDIANA PUTRI

NIM: 04221024

Diajukan guna memenuhi persyaratan
Untuk memperoleh gelar Sarjana Komputer (S.Kom)
Pada Program Studi Sistem Informasi
Fakultas Ilmu Komputer
Universitas Narotama Surabaya

Surabaya, 23 Juli 2025

Mengetahui/Menyetujui

Dosen Pembimbing,



Dr. Aryo Nugroho, S.T., S.Kom., M.T

NIDN: 0721077001

SURAT PERNYATAAN

Dengan ini saya menyatakan bahwa dalam Laporan Tugas Akhir ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan disuatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat Karya/Pendapat yang pernah ditulis oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam Daftar Acuan/Daftar Pustaka.

Apabila ditemukan suatu Jiplakan/Plagiat maka saya bersedia menerima akibat berupa sanksi Akademis dan sanksi lain yang diberikan oleh yang berwenang sesuai ketentuan peraturan dan perundang-undangan yang berlaku.

Surabaya, 7 Agustus 2025



Preity Deldiana Putri
NIM: 04221024

KATA PENGANTAR

Puji syukur kehadiran Tuhan Yang Maha Esa atas segala limpahan rahmat, hidayah dan karunia-Nya yang telah memberikan penulis kesehatan, kesempatan dan kemudahan sehingga dapat menyelesaikan penyusunan skripsi ini. Shalawat serta salam senantiasa tercurahkan kepada Nabi Muhammad SAW. Pada dasarnya, penelitian yang berjudul “Analisis Sentimen Produk Pembersih Wajah Menggunakan Metode Gaussian Naïve Bayes” merupakan bentuk pertanggungjawaban penulis untuk memenuhi salah satu syarat memperoleh gelar sarjana.

Penulis menyadari bahwa dalam penyusunan laporan ini tidak terlepas dari adanya dukungan, bantuan maupun arahan baik dalam hal pikiran, waktu dan tenaga. Maka dalam kesempatan ini penulis menyampaikan terimakasih kepada:

1. Allah SWT yang telah melimpahkan rahmat-Nya sehingga diberikan kesempatan, kemudahan dan kelancaran dalam melaksanakan penelitian hingga penyusunan skripsi ini.
2. Bapak Dr. Arasy Alimudin, SE, MM selaku Rektor Universitas Narotama Surabaya.
3. Bapak Eman Setiawan, S. Kom., M. M. selaku Ketua Program Studi Sistem Informasi.
4. Bapak Dr. Aryo Nugroho, S. T., S. Kom., M. T. selaku Dosen Pembimbing serta penguji yang telah memberikan arahan dan inspirasi mulai dari penelitian hingga penyusunan skripsi.

5. Seluruh dosen Fakultas Ilmu Komputer Universitas Narotama Surabaya yang telah membekali penulis dengan berbagai ilmu selama mengikuti perkuliahan sampai akhir penulisan skripsi.
6. Cinta pertama dan panutanku, Bapak Yon Zaujiyanto, terimakasih selalu berjuang untuk kehidupan penulis, beliau memang tidak sempat merasakan Pendidikan sampai bangku perkuliahan. Namun beliau mampu mendidik, memotivasi, memberikan dukungan hingga penulis mampu menyelesaikan studinya sampai sarjana. Sehat selalu dan panjang umur karena bapak harus selalu ada disetiap perjuangan dan pencapaian hidup penulis.
7. Pintu surgaku, Ibu Siti Nurhasanah, yang selalu menjadi penyemangat penulis. Yang tidak henti-hentinya memberikan kasih sayang dengan penuh cinta, memberikan motivasi, dukungan serta do'a dan nasihat hingga penulis mampu menyelesaikan studinya sampai sarjana. Sehat selalu dan panjang umur karena ibu harus selalu ada disetiap perjuangan dan pencapaian hidup penulis.
8. Kepada seseorang yang tak kalah penting kehadirannya, Roby Alansyah, S. H. Terimakasih telah membersamai dan menjadi bagian dalam perjalanan hidup penulis mulai dari awal masa SMA hingga penyusunan skripsi ini selesai. Berkontribusi banyak dalam penulisan karya tulis ini, baik tenaga, waktu maupun ide kepada penulis. Telah menjadi rumah, pendamping dalam segala hal yang menemani, mendukung maupun menghibur dalam kesedihan, mendengar keluh kesah, memberi semangat untuk pantang menyerah.
9. Teman seperjuangan Program Studi Sistem Informasi Angkatan 2021 yang selalu memberikan semangat, dukungan dan bantuan selama masa perkuliahan.

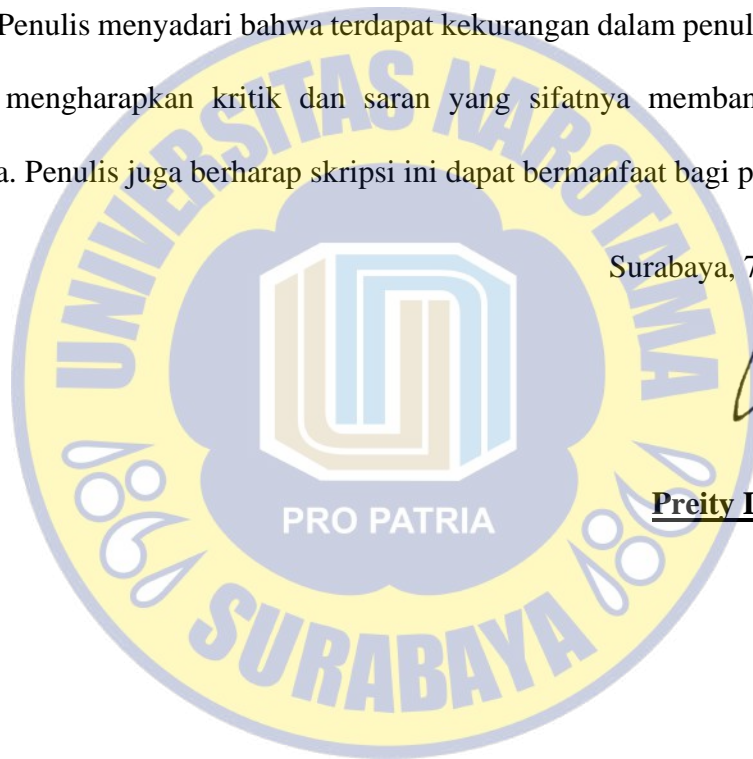
10. Terakhir, terimakasih untuk diri sendiri, Preity Deldiana Putri. Apresiasi sebesar-besarnya karena telah bertanggung jawab untuk menyelesaikan apa yang telah dimulai. Terimakasih karena terus berusaha dan tidak menyerah, serta senantiasa menikmati setiap prosesnya yang bisa dibilang tidak mudah.

Akhir kata penulis mengucapkan maaf yang sebesar-besarnya apabila terdapat kekurangan dan kesalahan baik dalam penulisan maupun penyusunan skripsi. Penulis menyadari bahwa terdapat kekurangan dalam penulisan skripsi ini, penulis mengharapkan kritik dan saran yang sifatnya membangun dari para pembaca. Penulis juga berharap skripsi ini dapat bermanfaat bagi pembaca.

Surabaya, 7 Agustus 2025



Preity Deldiana Putri



ABSTRAK

Pertumbuhan industri kecantikan di Indonesia mendorong meningkatnya kebutuhan konsumen terhadap informasi produk yang relevan dan terpercaya, khususnya melalui ulasan pengguna di platform daring seperti Female Daily. Penelitian ini bertujuan untuk menganalisis sentimen konsumen terhadap lima merek produk pembersih wajah (Acnes, Emina, Hadalabo, Senka, dan Wardah) dengan menggunakan metode *Gaussian Naïve Bayes* (GNB). Sebanyak 5.000 ulasan dikumpulkan melalui teknik *web scraping*, kemudian diproses menggunakan tahapan *cleaning*, *tokenizing*, *stopword*, dan *stemming*. Data selanjutnya diklasifikasikan ke dalam tiga kategori atribut produk: *skin type*, *function*, dan *texture*.

Hasil penelitian menunjukkan bahwa algoritma GNB mampu melakukan klasifikasi dengan tingkat akurasi tertinggi pada produk Acnes (95%) dan terendah pada Senka (76%). Kategori *skin type* merupakan atribut yang paling akurat teridentifikasi oleh model, sedangkan *texture* menunjukkan kecenderungan tertukar dengan *function* akibat kemiripan konteks. Visualisasi melalui TF-IDF dan *word cloud* juga mendukung hasil dari penelitian dengan memperlihatkan kata kunci dominan dalam setiap kategori. Penelitian ini memberikan kontribusi dalam pemanfaatan analisis sentimen untuk mendukung produsen memahami preferensi konsumen serta membantu konsumen dalam memilih produk sesuai kebutuhan kulit mereka.

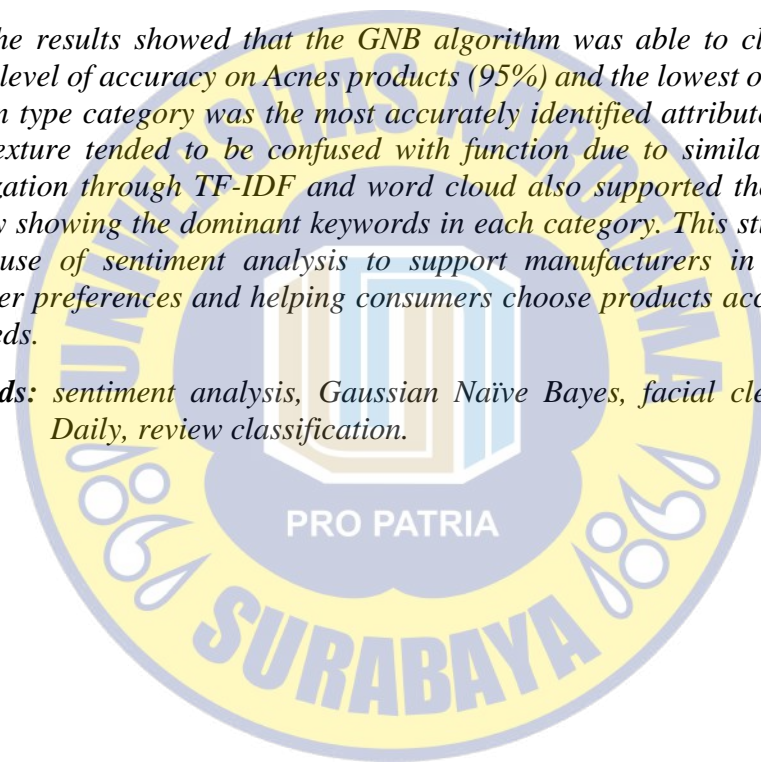
Kata kunci: analisis sentimen, Gaussian Naïve Bayes, pembersih wajah, Female Daily, klasifikasi ulasan.

ABSTRACT

The growth of the beauty industry in Indonesia has driven the increasing consumer demand for relevant and reliable product information, especially through user reviews on online platforms such as Female Daily. This study aims to analyze consumer sentiment towards five brands of facial cleansers (Acnes, Emina, Hadalabo, Senka, and Wardah) using the Gaussian Naïve Bayes (GNB) method. A total of 5,000 reviews were collected through web scraping techniques, then processed using the stages of cleaning, tokenizing, stopword, and stemming. The data was then classified into three product attribute categories: skin type, function, and texture.

The results showed that the GNB algorithm was able to classify with the highest level of accuracy on Acnes products (95%) and the lowest on Senka (76%). The skin type category was the most accurately identified attribute by the model, while texture tended to be confused with function due to similarity in context. Visualization through TF-IDF and word cloud also supported the results of the study by showing the dominant keywords in each category. This study contributes to the use of sentiment analysis to support manufacturers in understanding consumer preferences and helping consumers choose products according to their skin needs.

Keywords: *sentiment analysis, Gaussian Naïve Bayes, facial cleanser, Female Daily, review classification.*



DAFTAR ISI

LEMBAR PERSETUJUAN.....	iii
LEMBAR PENGESAHAN.....	iii
SURAT PERNYATAAN.....	iv
MOTTO DAN PERSEMBAHAN.....	vi
KATA PENGANTAR.....	vii
ABSTRAK.....	x
<i>ABSTRACT</i>	xi
DAFTAR ISI.....	xii
DAFTAR GAMBAR.....	xv
DAFTAR TABEL.....	xvi
DAFTAR LAMPIRAN.....	xvii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	5
1.3 Batasan Masalah.....	5
1.4 Tujuan Penelitian.....	6
1.5 Manfaat Penelitian.....	6
BAB II TINJAUAN PUSTAKA.....	8
2.1 Definisi & Teori.....	8
2.1.1 <i>Natural Language Processing</i>	8
2.1.2 <i>Text Mining</i>	9
2.1.3 Analisis Sentimen.....	9

2.1.4 Female Daily.....	11
2.1.5 Pembersih Wajah/ <i>Facial Wash</i>	13
2.1.6 Himpunan.....	13
2.1.7 Probabilitas	14
2.1.8 <i>Naïve Bayes</i>	16
2.1.9 <i>Gaussian Naïve Bayes</i>	17
2.2 Penelitian Terdahulu.....	17
BAB III METODE PENELITIAN.....	20
3.1 Jenis Penelitian.....	20
3.2 Desain Penelitian	20
3.3 Sumber Data.....	22
3.4 Teknik Pengumpulan Data	22
3.5 Instrumen Penelitian.....	23
3.6 Analisis Data	26
BAB IV HASIL DAN PEMBAHASAN.....	28
4.1 Tahap Pengambilan data.....	28
4.2 Tahap <i>Preprocessing</i>	30
4.3 Pemodelan	37
4.4 Evaluasi	38
4.4.1 <i>Term Frequency-Inverst Document Frequency</i>	38
4.4.2 <i>Confusion matrix</i>	49
4.4.3 Akurasi.....	62
4.5 Word Cloud.....	64
BAB V KESIMPULAN DAN SARAN.....	78
5.1 Kesimpulan	78

5.2 Saran	79
DAFTAR PUSTAKA	81
LAMPIRAN	85



DAFTAR GAMBAR

Gambar 2. 1 Logo Female Daily	11
Gambar 2. 2 Website Female Daily	12
Gambar 3. 1 Desain Penelitian	21
Gambar 4. 1 Ulasan Pelanggan	29
Gambar 4. 2 Tahapan WebScrapping	29
Gambar 4. 3 Data Ulasan	30
Gambar 4. 4 Source Code Preprocessing	32
Gambar 4. 5 Source Code Pemodelan	37
Gambar 4. 6 Source Code TF-IDF	40
Gambar 4. 7 TF-IDF Acnes	42
Gambar 4. 8 TF-IDF Emina	44
Gambar 4. 9 TF-IDF Hadalabo	45
Gambar 4. 10 TF-IDF Senka	47
Gambar 4. 11 TF-IDF Wardah	48
Gambar 4. 12 Source Code <i>Confusion matrix</i>	50
Gambar 4. 13 Confusion Matrik Acnes	53
Gambar 4. 14 Confusion Matrik Emina	55
Gambar 4. 15 Confusion Matrik Hadalabo	57
Gambar 4. 16 Confusion Matrik Senka	59
Gambar 4. 17 Confusion Matrik Wardah	61
Gambar 4. 18 Source Code Akurasi	63
Gambar 4. 19 Akurasi	64
Gambar 4. 20 Source Code Word Cloud	65
Gambar 4. 21 Word Cloud Acnes	67
Gambar 4. 22 Word Cloud Emina	69
Gambar 4. 23 Word Cloud Hadalabo	72
Gambar 4. 24 Word Cloud Senka	74
Gambar 4. 25 Word Cloud Wardah	75

DAFTAR TABEL

Tabel 2. 1 Penelitian Terdahulu	19
Tabel 3. 1 Tools.....	25
Tabel 4. 1 Cleaning	33
Tabel 4. 2 Tokenizing	34
Tabel 4. 3 Stopword	35
Tabel 4. 4 Stemming	36
Tabel 4. 5 Penjelasan Warna Confusion matrix.....	50
Tabel 4. 6 Confusion matrix Acnes.....	51
Tabel 4. 7 Confusion matrix Emina	53
Tabel 4. 8 Confusion matrix Hadalabo	55
Tabel 4. 9 Confusion matrix Senka.....	57
Tabel 4. 10 Confusion matrix Wardah.....	59
Tabel 4. 11 Frekuensi Word Cloud Acnes	66
Tabel 4. 12 Frekuensi Word Cloud Emina.....	68
Tabel 4. 13 Frekuensi Word Cloud Hadalabo.....	70
Tabel 4. 14 Frekuensi Word Cloud Senka	72
Tabel 4. 15 Frekuensi Word Cloud Wardah	74
Tabel 4. 16 Rekomendasi Produk	76