PERSEMBAHAN

"Sesungguhnya sesudah kesulitan itu ada kemudahan. Maka apabila kamu telah selesai (dari suatu urusan), kerjakanlah dengan sungguh-sungguh (urusan) yang lain." (Q.S. Al-Insyirah: 6-7)

Dengan mengucap puji syukur kepada Allah S.W.T, skripsi ini kupersembahkan untuk:

1. Kedua Orang Tua Tercinta, terima kasih yang tak terhingga atas doa yang tak pernah putus di setiap sepertiga malam, tetesan keringat, pengorbanan, dan kasih sayang tanpa batas. Skripsi ini adalah bukti kecil dari perjuangan panjang kalian dalam mendidik dan membesarkan saya. Semoga Allah SWT senantiasa menjaga dan merahmati kalian di dunia dan akhirat.
2. Dosen Pembimbing Bp. Ruhul Amin, M.Kom. Terima kasih atas ilmu, kesabaran, arahan, dan bimbingannya selama proses penyusunan skripsi ini. Semoga setiap ilmu yang ditularkan menjadi amal jariyah yang mengalir tiada putus.
3. Diri Saya Sendiri karena telah bertahan, tidak menyerah meski berkali-kali dihampiri rasa lelah dan lewati proses rumit pemrosesan data serta pemodelan. Kamu hebat telah melangkah sejauh ini
4. Kepada sesorang yang tidak bisa saya sebutkan Namanya yang telah ikut andil dan mensuport saya dalam penyusunan Skripsi ini.

Tanpa mereka,

Aku dan karya ini tidak akan pernah ada

UNIVERSITAS NUSA MANDIRI

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH

Yang bertanda tangan dibawah ini, saya :

Nama : Andi Tyas Saputra

NIM : 12250094

Program Studi : Informatika

Fakultas : Teknologi Informasi

Perguruan tinggi : Universitas Nusa Mandiri

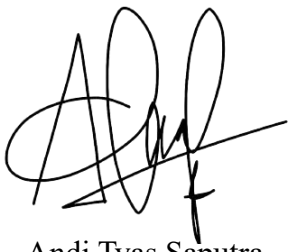
Dengan ini menyetujui untuk memberikan ijin kepada pihak Universitas Nusa Mandiri, Hak Bebas Royalti Non-Eksklusif (Non- exclusive Royalti-Free Right) atas karya ilmiah kami yang berjudul: “Judul Skripsi”, beserta perangkat yang diperlukan (apabila ada).

Dengan Hak Bebas Royalti Non-Eksklusif ini kepada pihak Universitas Nusa Mandiri berhak menyimpan, mengalih-media atau format-kan, mengelolaannya dalam pangkalan data (database), mendistribusikannya dan menampilkan atau mempublikasikannya di internetatau media lain untuk kepentingan akademis tanpa perlu meminta ijin dari kami selama tetap mencantumkan nama kami sebagai penulis/pencipta karya ilmiah tersebut.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Universitas Nusa Mandiri, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini saya buat dengan sebenarya.

Dibuat : Jakarta
Pada Tanggal : 26 Juli 2026
Yang Menyatakan,



Andi Tyas Saputra

UNIVERSITAS NUSA MANDIRI

PERSETUJUAN DAN PENGESAHAN TUGAS AKHIR

Tugas Akhir ini diajukan oleh:

Nama	:	Andi Tyas Saputra
NIM	:	12250094
Program Studi	:	Informatika
Fakultas	:	Teknologi Informasi
Jenjang	:	Strata Satu (S1)
Judul Tugas Akhir	:	Analisis Sentimen Review Skincare di Marketplace Tokopedia Menggunakan Algoritma K-Nearest Neighbor

Telah dipertahankan pada periode 2026-1 dihadapan penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh Sarjana Komputer (S.Kom) pada Program Sarjana Program Studi Informatika Fakultas Teknologi Informasi di Universitas Nusa Mandiri.

Jakarta, 18 Agustus 2026

PEMBIMBING TUGAS AKHIR

Dosen Pembimbing : Dr. Ruhul Amin, M.Kom.

Ditandatangani Digital:
Ruhul Amin
2026-08-19 04:34:33

.....

D E W A N P E N G U J I

Penguji I : Frisma Handayanna, M.Kom

Ditandatangani Digital:
Frisma Handayanna
2026-08-18 11:52:18

.....

Penguji II : Sita Anggraeni , M.Kom.

Ditandatangani Digital:
Sita Anggraeni
2026-08-18 10:43:35

.....

LEMBAR PEDOMAN PENGGUNAAN HAK CIPTA

Skripsi yang berjudul “Judul Skripsi” adalah hasil karya tulis asli Nama Mahasiswa dan bukan hasil terbitan sehingga peredaran karya tulis hanya berlaku di lingkungan akademik saja, serta memiliki hak cipta. Oleh karena itu, dilarang keras untuk menggandakan baik sebagian maupun seluruhnya karya tulis ini, tanpa seizin penulis.

Referensi kepustakaan diperkenankan untuk dicatat tetapi pengutipan atau peringkasan isi tulisan hanya dapat dilakukan dengan seizin penulis dan disertai ketentuan pengutipan secara ilmiah dengan menyebutkan sumbernya.

Untuk keperluan perizinan pada pemilik dapat menghubungi informasi yang tertera di bawah ini:

Nama	: Andi Tyas Saputra
Alamat	: Jalan. Rukun 1 no.24, Kel. Pekayon, Kec. Pasar Rebo, Kota Jakarta timur
No. Telp	: 085693254307
E-mail	: andtsputra@gmail.com

UNIVERSITAS
NUSA MANDIRI

KATA PENGANTAR

Alhamdulillah, dengan mengucapkan puji syukur kehadirat Allah SWT, yang telah melimpahkan rahmat dan karunia-Nya, sehingga pada akhirnya penulis dapat menyelesaikan tugas ini dengan baik. Dimana Skripsi ini penulis sajikan dalam bentuk buku yang sederhana. Adapun judul Skripsi, yang penulis ambil sebagai berikut,

"Analisis Sentimen *Review Skincare* Di Marketplace Tokopedia Menggunakan Algoritma *K-nearest Neighbor*"

Tujuan penulisan Skripsi ini dibuat sebagai salah satu syarat kelulusan Program Sarjana Universitas Nusa Mandiri. Sebagai bahan penulisan diambil berdasarkan hasil penelitian (eksperimen), observasi dan beberapa sumber literatur yang mendukung penulisan ini. Penulis menyadari bahwa tanpa bimbingan dan dorongan dari semua pihak, maka penulisan Skripsi ini tidak akan lancar. Oleh karena itu pada kesempatan ini, izinkanlah penulis menyampaikan ucapan terima kasihkepada:

1. Rektor Universitas Nusa Mandiri.
2. Wakil Rektor I Bidang Akademik Universitas Nusa Mandiri.
3. Dekan Fakultas Teknologi Informasi Universitas Nusa Mandiri.
4. Ketua Program Studi Informatika Universitas Nusa Mandiri.
5. Bapak Ruhul Amin, M.Kom. selaku Dosen Pembimbing Skripsi.
6. Bapak/ibu dosen Program Studi Informatika Universitas Nusa Mandiri yang telah memberikan penulis dengan semua bahan yang diperlukan.
7. Staff / karyawan / dosen di lingkungan Fakultas Teknologi Informasi Universitas Nusa Mandiri.

UNIVERSITAS NUSA MANDIRI

8. Orang tua tercinta yang telah memberikan dukungan moral maupun spritual.
9. Rekan-rekan mahasiswa kelas 12.8C.10

Serta semua pihak yang terlalu banyak untuk disebut satu persatu sehingga terwujudnya penulisan ini. Penulis menyadari bahwa penulisan Skripsi ini masih jauh sekali dari sempurna, untuk itu penulis mohon kritik dan saran yang bersifat membangun demi kesempurnaan penulisan dimasa yang akan datang.

Akhir kata semoga Skripsi ini dapat berguna bagi penulis khususnya dan bagi para pembaca yang berminat pada umumnya.

Jakarta, 26 Juli 2026

Penulis



Andi Tyas Saputra

UNIVERSITAS
NUSA MANDIRI

ABSTRAK

Andi Tyas Saputra (12250094) "Analisis Sentimen *Review Skincare* Di Marketplace Tokopedia Menggunakan Algoritma *K-nearest Neighbor*"

Permasalahan dalam penelitian ini adalah besarnya volume ulasan produk skincare di Tokopedia yang bersifat tidak terstruktur dan menggunakan bahasa tidak baku, sehingga pembeli maupun penjual kesulitan dalam mengekstrak informasi sentimen konsumen secara efisien. Tujuan penelitian ini adalah mengklasifikasikan sentimen ulasan produk skincare (Skintific) di Tokopedia ke dalam kategori positif, netral, dan negatif, mengukur kinerja pengklasifikasi, serta membangun aplikasi prediksi berbasis web. Metode yang diusulkan memanfaatkan algoritma K-Nearest Neighbor (KNN) dengan pembobotan fitur Term Frequency-Inverse Document Frequency (TF-IDF) dan kalkulasi Euclidean Distance (K=5). Tahapan penelitian meliputi scraping 2.000 data ulasan (diperoleh 1.911 ulasan bersih), pra-pemrosesan teks (case folding, cleaning, tokenizing, stopword removal, dan stemming), pembagian data rasio 80:20 (1.528 data latih dan 383 data uji), serta implementasi antarmuka interaktif menggunakan framework Streamlit. Hasil yang diperoleh menunjukkan bahwa model KNN mencapai akurasi global sebesar 95,56% (366 dari 383 data uji terprediksi tepat). Performa pada kelas mayoritas (positif) sangat optimal dengan presisi 96%, recall 100%, dan F1-score 98%. Kendati terdapat keterbatasan klasifikasi pada kelas minoritas (netral dan negatif) akibat ketidakseimbangan data (data imbalance), aplikasi web Streamlit berhasil memfasilitasi visualisasi serta prediksi sentimen ulasan secara interaktif untuk pengujian teks tunggal dan berkas skala besar.

Kata Kunci : Analisis Sentimen, *K-Nearest Neighbor*, *Skincare*, Tokopedia, TF-IDF, *Streamlit*.

UNIVERSITAS NUSA MANDIRI

ABSTRACT

Andi Tyas Saputra (12250094) "Sentiment Analysis of Skincare Product Reviews on Tokopedia Marketplace Using the K-Nearest Neighbor Algorithm"

The research problem stems from the massive volume of unstructured skincare product reviews on Tokopedia containing informal language, which makes it difficult for consumers and sellers to extract public sentiment efficiently. The research objective is to classify skincare review sentiment (Skintific) on Tokopedia into positive, neutral, and negative categories, evaluate model performance, and deploy an interactive web-based prediction application. The proposed method utilizes the K-Nearest Neighbor (KNN) algorithm combined with Term Frequency-Inverse Document Frequency (TF-IDF) feature weighting and Euclidean Distance (K=5). The research steps encompass scraping 2,000 raw reviews (yielding 1,911 clean data points), text preprocessing (case folding, cleaning, tokenization, stopword removal, and stemming), an 80:20 data split (1,528 training data and 383 testing data), and interface implementation using the Streamlit framework. The obtained results demonstrate that the KNN model achieved a global accuracy of 95.56% (366 of 383 test samples correctly classified). The performance for the majority class (positive) was optimal, reaching 96% precision, 100% recall, and a 98% F1-score. Although classification on minority classes (neutral and negative) was limited due to extreme data imbalance, the Streamlit web application successfully facilitates interactive review sentiment prediction and visualization for single-text inputs and bulk file testing.

Keywords: Sentiment Analysis, K-Nearest Neighbor, Skincare, Tokopedia, TF-IDF, Streamlit.

UNIVERSITAS
NUSA MANDIRI

DAFTAR ISI

LEMBAR JUDUL i
PERSEMBAHAN ii
LEMBAR PERNYATAAN KEASLIAN SKRIPSI iii
LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH iv
LEMBAR PERSETUJUAN DAN PENGESAHAN SKRIPSI v
LEMBAR PEDOMAN PENGGUNAAN HAK CIPTA vi
KATA PENGANTAR vii
ABSTRAK ix
ABSTRACT x
DAFTAR ISI xi
DAFTAR GAMBAR xiii
DAFTAR TABEL xiv
DAFTAR LAMPIRAN xv
BAB I PENDAHULUAN 1
1.1 Latar Belakang Masalah 1
1.2 Identifikasi Permasalahan 3
1.3 Perumusan Masalah 3
1.4 Maksud dan Tujuan 4
1.5 Metodologi Penelitian 5
1.5.1 Teknik Pengumpulan Data 5
1.6 Ruang Lingkup 6
BAB II LANDASAN TEORI 7
2.1 Tinjauan Pustaka 7
2.1.1. Analisis Sentimen 7
2.1.2. Text Mining 7
2.1.3. K-Nearest Neighbor 8
2.1.4. Pengertian Machine Learning 8
2.2 Penelitian Terkait 8
BAB III ANALISA SISTEM BERJALAN 14
3.1 Alur Penelitian 14
3.2 Pengumpulan Data 15

UNIVERSITAS NUSA MANDIRI

3.2.1 Sumber Data 15
3.2.2 Statistik Deskriptif Data 16
3.4 Perancangan model 21
3.5 Skenario Pengujian 24
BAB IV RANCANGAN APLIKASI USULAN 30
4.1 Hasil Penelitian 30
4.1.1. Scraping Data 30
4.1.2. Labelling Data 31
4.1.3. Text Preprocessing 31
4.1.4. Splitting Data 32
4.1.5. Term Frequency-Inverse Document Frequency (TF-IDF) 32
4.1.6. K-Nearest Neighbor 33
4.1.7. Evaluasi Model (Confusion Matrix) 33
4.1.8. Visualisasi 35
4.2 Hasil Implementasi Antarmuka 35
4.2.1. Halaman Dashboard Dan Ringkasan (Beranda Utama) 36
4.2.2. Halaman Pembobotan TF-IDF 37
4.2.3. Halaman Prediksi Teks Tunggal (Uji Interaktif) 38
4.2.4. Halaman Prediksi Skala Besar 39
4.2.5. Halaman Eksplorasi Dataset 40
BAB V PENUTUP 41
5.1 Kesimpulan 41
5.2 Saran 42
DAFTAR PUSTAKA 43
DAFTAR RIWAYAT HIDUP 45
LEMBAR BIMBINGAN SKRIPSI 46
LAMPIRAN 47
LAMPIRAN. A Bukti Hasil Pengecekan Plagiarisme 47
LAMPIRAN. B Publish Artikel Ilmiah 48
UNIVERSITAS NUSA MANDIRI

DAFTAR GAMBAR

Gambar III. 1 Kerangka Penelitian 14
Gambar VI. 1 *Confusion Matrix* 34
Gambar VI. 2 Hasil Distribusi Sentimen Berdasarkan Label 35
Gambar VI. 3 Halaman Dashboard dan ringkasan 36
Gambar VI. 4 Halaman Dashboard dan ringkasan 37
Gambar VI. 5 Halaman Dashboard dan ringkasan 37
Gambar VI. 6 Halaman Pembobotan TF-IDF 38
Gambar VI. 7 Halaman Pembobotan TF-IDF 38
Gambar VI. 8 Halaman Prediksi Teks Tunggal 39
Gambar VI. 9 Halaman Prediksi Skala Besar 39
Gambar VI. 10 Halaman Eksplorasi Dataset 40

UNIVERSITAS
NUSA MANDIRI

DAFTAR TABEL

Tabel III. 1 Contoh Hasil *Labeling* 15
Tabel III. 2 Deskriptif Data 17
Tabel III. 3 *Confusion Matrix* 26
Tabel IV. 1 Contoh Hasil dari *Scraping* 30
Tabel IV. 2 Hasil Dari *Labelling* Data 31
Tabel IV. 3 Hasil Dari *Text Preprocessing* 32
Tabel IV. 4 *Classification Report* 33

UNIVERSITAS
NUSA MANDIRI

DAFTAR LAMPIRAN

LAMPIRAN. A Bukti Hasil Pengecekan Plagiarisme 47
LAMPIRAN. B Publish Artikel Ilmiah..... 48

UNIVERSITAS
NUSA MANDIRI