

TUGAS AKHIR

PENERAPAN ALGORITMA *MULTINOMIAL NAÏVE BAYES* DALAM KLASIFIKASI SENTIMEN PENGGUNA SHOPEE TERHADAP PRODUK *FACIAL WASH* KAHF

Diajukan sebagai salah satu syarat untuk kelulusan
Strata Satu (S-1) Prodi Teknik Informatika Fakultas Teknik
Universitas Muhammadiyah Jember



Oleh:
Farendika Rezzi
NIM. 2110651061

**UNIVERSITAS MUHAMMADIYAH JEMBER
FAKULTAS TEKNIK
PROGRAM STUDI TEKNIK INFORMATIKA
2025**

LEMBAR PERSETUJUAN SIDANG TUGAS AKHIR

**PENERAPAN ALGORITMA *MULTINOMIAL NAÏVE BAYES* DALAM
KLASIFIKASI SENTIMEN PENGGUNA SHOPEE TERHADAP PRODUK
FACIAL WASH KAHF**

Oleh :

Farendika Rezzi

2110651061

Telah disetujui bahwa Laporan Tugas Akhir ini untuk diajukan pada sidang Tugas Akhir sebagai salah satu syarat kelulusan dan mendapatkan gelas Sarjana Komputer (S.Kom)

Di

Universitas Muhammadiyah Jember

Disetujui oleh,

Pembimbing I

Pembimbing II


Ulya Anisatur Rosyidah, M.Kom
NIDN. 0714078704


Ginanjar Abdurrahman, S.Si., M.Pd.
NIDN. 0710037903

LEMBAR PENGESAHAN
PENERAPAN ALGORITMA MULTINOMIAL NAÏVE BAYES DALAM
KLASIFIKASI SENTIMEN PENGGUNA SHOPEE TERHADAP PRODUK
FACIAL WASH KAHF

Oleh:

Farendika Rezzi

2110651061

Telah mempertanggungjawabkan Laporan Tugas Akhirnya pada sidang Tugas Akhir tanggal sebagai salah satu syarat kelulusan dan mendapatkan gelar Sarjana Komputer (S.Kom)

Di

Universitas Muhammadiyah Jember

Disetujui oleh,

Dosen Penguji I

Dosen Pembimbing I



Henny Wahyu Sulistyo, M.Kom
NIDN. 0728079101



Ulva Anisatur Rosyidah, M.Kom
NIDN. 0714078704

Dosen Penguji II

Dosen Pembimbing II



Qurrota A'yun M.Pd
NIDN. 0703069002



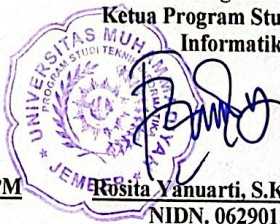
Ginanjar Abdurrahman, S.Si., M.Pd.
NIDN. 0710037903

**Mengesahkan,
Dekan Fakultas Teknik**

**Mengetahui,
Ketua Program Studi Teknik
Informatika**



Dr. Ir. Muhtar, ST., MT., IPM
NIDN. 0014067301



Kusita Yanuarti, S.Kom., M.Cs
NIDN. 0629018601

LEMBAR PERNYATAAN

Yang bertanda tangan di bawah ini:

Nama : Farendika Rezzi

NIM : 2110651061

Dengan ini saya menyatakan dengan sebenar-benarnya bahwa Tugas Akhir berjudul "*Penerapan Algoritma Multinomial Naïve Bayes dalam Klasifikasi Sentimen Pengguna Shopee terhadap Produk Facial Wash Kahf*" sepenuhnya merupakan hasil karya saya sendiri. Segala bentuk kutipan atau bagian yang berasal dari karya orang lain telah diberi tanda sitasi dan dicantumkan dalam daftar pustaka. Apabila di kemudian hari terbukti pernyataan ini tidak benar atau ditemukan pelanggaran terkait karya Tugas Akhir ini, saya bersedia menerima sanksi sesuai ketentuan akademik yang berlaku.

Jember, 15 Agustus 2025

Yang membuat pernyataan

Farendika Rezzi

NIM. 2110651061



KATA PENGANTAR

Puji syukur kehadiran Allah SWT atas rahmat, taufik, dan hidayah-Nya sehingga penulis dapat menyelesaikan skripsi berjudul "*Penerapan Algoritma Multinomial Naïve Bayes dalam Klasifikasi Sentimen Pengguna Shopee terhadap Produk Facial Wash Kahf*" sebagai salah satu syarat memperoleh gelar Sarjana Teknik pada Program Studi Teknik Informatika, Fakultas Teknik, Universitas Muhammadiyah Jember. Ucapan terima kasih penulis sampaikan kepada kedua orang tua tercinta atas doa dan dukungan yang tiada henti; Ibu Ulya Anisatur Rosyidah, M.Kom. selaku pembimbing I; Bapak Ginanjar Abdurrahman, S.Si., M.Pd. selaku pembimbing II; Bapak Henny Wahyu Sulisty, S.Kom., M.Kom. selaku dosen penguji I; Ibu Qurrota A'yun, M.Pd. selaku dosen penguji II; Ibu Rosita Yanuarti, S.Kom., M.Cs. selaku Ketua Program Studi; seluruh dosen dan staf yang telah memberikan ilmu serta bimbingan; serta teman-teman seperjuangan atas semangat dan kebersamaan yang berarti.

Penulis menyadari skripsi ini masih jauh dari sempurna, sehingga kritik dan saran yang membangun sangat diharapkan demi perbaikan di masa mendatang. Semoga karya ini bermanfaat bagi pembaca dan menjadi referensi bagi penelitian di bidang analisis sentimen dan *machine learning*.

Jember, 1 Agustus 2025


Farendika Rezzi

DAFTAR ISI

TUGAS AKHIR	i
LEMBAR PERSETUJUAN SIDANG TUGAS AKHIR	ii
LEMBAR PENGESAHAN	iii
LEMBAR PERNYATAAN	iv
KATA PENGANTAR.....	v
PERSEMBAHAN.....	vi
MOTTO	viii
DAFTAR ISI.....	ix
DAFTAR GAMBAR.....	xii
DAFTAR TABEL	xiii
ABSTRAK	xv
ABSTRACT	xvi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Tujuan Penelitian	4
1.4 Manfaat Penelitian	5
1.5 Batasan Penelitian	5
BAB II TINJAUAN PUSTAKA.....	7
2.1 Penelitian Terdahulu	7
2.2 Shopee	9
2.3 Kahf.....	11
2.4 <i>Text Mining</i>	11
2.5 <i>Preprocessing</i>	12
2.6 Pembobotan <i>TF-IDF</i>	13
2.7 <i>K-Fold Cross Validation</i>	15
2.8 <i>Stratified K-Fold Cross Validation</i>	16
2.9 <i>Synthetic Minority Over-sampling Technique (SMOTE)</i>	16
2.10 <i>Multinomial Naïve Bayes</i>	17
2.11 <i>Confusion Matrix</i>	18
2.12 <i>Python</i>	20

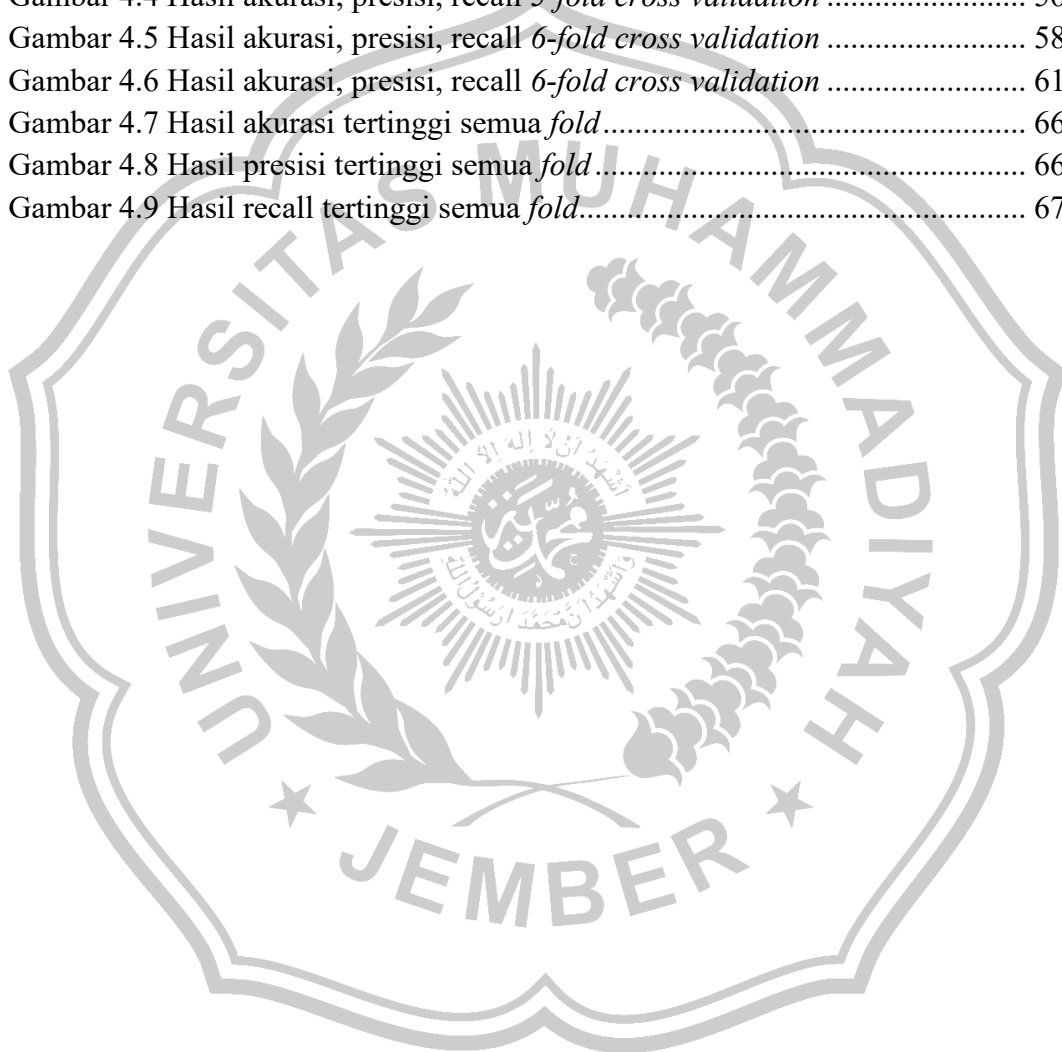
BAB III METODOLOGI PENELITIAN	22
3.1 Tahap Penelitian	22
3.2 Pengumpulan Data	22
3.3 Labelling Data	23
3.4 Preprocessing Text	23
3.5 Ekstraksi Fitur <i>TF-IDF</i>	27
3.5.1 Perhitungan <i>Term Frequency (TF)</i>	27
3.5.2 Perhitungan <i>Inverse Document Frequency (IDF)</i>	28
3.5.3 Perhitungan <i>TF-IDF</i>	30
3.6 Pembagian Data (<i>K-Fold Cross Validation</i>)	31
3.7 Penyeimbangan Data (<i>SMOTE</i>)	35
3.8 Klasifikasi <i>Multinomial Naïve Bayes</i>	35
3.9 Evaluasi Model	38
BAB IV HASIL DAN PEMBAHASAN	40
4.1 Pengumpulan Data	40
4.2 Labelling Data	40
4.3 Hasil <i>Preprocessing</i>	41
4.3.1 <i>Cleansing</i>	41
4.3.2 <i>Case Folding</i>	42
4.3.3 <i>Normalize</i>	43
4.3.4 <i>Tokenizing</i>	44
4.3.5 <i>Stopword Removal</i>	45
4.3.6 <i>Stemming</i>	46
4.4 Ekstraksi Fitur <i>TF-IDF</i>	47
4.5 Penyeimbangan Data (<i>SMOTE</i>)	49
4.5.1 <i>SMOTE Data Latih 4-Fold</i>	49
4.5.2 <i>SMOTE Data Latih 5-Fold</i>	50
4.5.3 <i>SMOTE Data Latih 6-Fold</i>	50
4.5.4 <i>SMOTE Data Latih 10-Fold</i>	51
4.6 Pembagian dan Pengujian Data	52
4.6.1 <i>4-Fold Cross Validation</i>	53
4.6.2 <i>5-Fold Cross Validation</i>	55
4.6.3 <i>6-Fold Cross Validation</i>	57
4.6.4 <i>10-Fold Cross Validation</i>	59

4.7 Evaluasi Model.....	66
BAB V KESIMPULAN DAN SARAN	68
5.1 Kesimpulan	68
5.2 Saran.....	68
DAFTAR PUSTAKA.....	70
LAMPIRAN.....	73



DAFTAR GAMBAR

Gambar 1.1 Grafik berdasarkan data riset internal tim compas.....	1
Gambar 2.1 <i>K-Fold Cross Validation</i>	15
Gambar 3.1 Alur Tahapan Penelitian.....	22
Gambar 4.1 Hasil <i>Scraping Data</i>	40
Gambar 4.2 Hasil <i>Labelling</i>	41
Gambar 4.3 Hasil akurasi, presisi, recall <i>4-fold cross validation</i>	54
Gambar 4.4 Hasil akurasi, presisi, recall <i>5-fold cross validation</i>	56
Gambar 4.5 Hasil akurasi, presisi, recall <i>6-fold cross validation</i>	58
Gambar 4.6 Hasil akurasi, presisi, recall <i>6-fold cross validation</i>	61
Gambar 4.7 Hasil akurasi tertinggi semua <i>fold</i>	66
Gambar 4.8 Hasil presisi tertinggi semua <i>fold</i>	66
Gambar 4.9 Hasil recall tertinggi semua <i>fold</i>	67



DAFTAR TABEL

Tabel 2.1 Penelitian Terdahulu	7
Tabel 2.2 <i>Confusion Matrix</i>	19
Tabel 3.1 Contoh Hasil <i>Scrapping</i>	22
Tabel 3.2 <i>Labelling Data</i>	23
Tabel 3.3 Tahap <i>Cleansing</i>	24
Tabel 3.4 Tahap <i>Case Folding</i>	24
Tabel 3.5 Tahap <i>Normalize</i>	25
Tabel 3.6 Tahap <i>Tokenizing</i>	25
Tabel 3.7 Tahap <i>Stopword removal</i>	26
Tabel 3.8 Tahap <i>Stemming</i>	26
Tabel 3.9 Dokumen untuk <i>TF-IDF</i>	27
Tabel 3.10 Perhitungan <i>TF</i>	27
Tabel 3.11 Perhitungan <i>IDF</i>	29
Tabel 3.12 Hasil pembobotan <i>TF-IDF</i>	30
Tabel 3.13 Probabilitas <i>Term</i> pada Kelas Data Uji	36
Tabel 3.14 Hasil Klasifikasi <i>Multinomial Naïve Bayes</i>	37
Tabel 3.15 <i>Confusion Matrix</i>	38
Tabel 3.16 Hasil Kriteria	38
Tabel 4.1 Hasil <i>Cleansing</i>	41
Tabel 4.2 Hasil <i>Case Folding</i>	42
Tabel 4.3 Hasil <i>Normalize</i>	43
Tabel 4.4 Hasil <i>Tokenizing</i>	44
Tabel 4.5 Hasil <i>Stopword Removal</i>	45
Tabel 4.6 Hasil <i>Stemming</i>	47
Tabel 4.7 <i>Ekstraksi Fitur</i> (Pembobotan <i>TF-IDF</i>)	48
Tabel 4.8 <i>SMOTE</i> Data Latih 4-Fold	50
Tabel 4.9 <i>SMOTE</i> Data Latih 5-Fold	50
Tabel 4.10 <i>SMOTE</i> Data Latih 6-Fold	51
Tabel 4.11 <i>SMOTE</i> Data Latih 10-Fold	51
Tabel 4.12 Hasil <i>confusion matrix 4-fold</i> tiap langkah	53
Tabel 4.13 Hasil akurasi, presisi, recall 4-fold <i>cross validation</i>	54
Tabel 4.14 <i>Confusion matrix 4-fold cross validation</i>	54
Tabel 4.15 Hasil <i>confusion matrix 5-fold</i> tiap langkah	55
Tabel 4.16 Hasil akurasi, presisi, recall 6-fold <i>cross validation</i>	56
Tabel 4.17 <i>Confusion matrix 5-fold cross validation</i>	56
Tabel 4.18 Hasil <i>confusion matrix 6-fold</i> tiap langkah	57
Tabel 4.19 Hasil akurasi, presisi, recall 6-fold <i>cross validation</i>	58
Tabel 4.20 <i>Confusion matrix 6-fold cross validation</i>	59
Tabel 4.21 Hasil <i>confusion matrix 10-fold</i> tiap langkah	59
Tabel 4.22 Hasil akurasi, presisi, recall 10-fold <i>cross validation</i>	60

Tabel 4.23 <i>Confusion matrix 10-fold cross validation</i>	61
Tabel 4.24 Rekapitulasi hasil akurasi, <i>preisi</i> , <i>recall</i>	62
Tabel 4.25 Contoh Komentar Uji dan Hasil Prediksi Sentimen pada Langkah ke-6 (<i>10-Fold Cross Validation</i>).....	63
Tabel 4.26 Hasil sentimen <i>aktual</i> dan <i>prediksi 10-fold</i> langkah ke-6.....	63
Tabel 4.27 <i>Confusion matrix 10-fold cross validation langkah uji 4</i>	65

