

**PENERAPAN ALGORITMA MULTINOMIAL NAÏVE BAYES
UNTUK KLASIFIKASI SENTIMEN PADA HASHTAG
#KABURAJADULU DI MEDIA SOSIAL X**



Muhammad Ni'mad Rahmatullah

2110651103

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS TEKNIK

UNIVERSITAS MUHAMMADIYAH JEMBER

2026

TUGAS AKHIR

PENERAPAN ALGORITMA MULTINOMIAL NAÏVE BAYES UNTUK KLASIFIKASI SENTIMEN PADA HASHTAG #KABURAJADULU DI MEDIA SOSIAL X

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



Muhammad Ni'mad Rahmatullah

2110651103

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS TEKNIK

UNIVERSITAS MUHAMMADIYAH JEMBER

2026

PERNYATAAN ORISINALITAS

Saya yang bertanda tangan di bawah ini:

Nama : Muhammad Ni'mad Rahmatullah

NIM : 2110651103

Program Studi : Teknik Informatika

Judul Tugas Akhir : PENERAPAN ALGORITMA MULTINOMIAL NAÏVE
BAYES UNTUK KLASIFIKASI SENTIMEN PADA
HASHTAG #KABURAJADULU DI MEDIA SOSIAL X

Dengan ini menyatakan bahwa karya ilmiah yang saya susun adalah murni hasil pemikiran, penelitian, dan penyusunan saya sendiri. Segala sumber informasi yang berasal atau dikutip dari karya orang lain, baik berupa publikasi, artikel, buku, maupun sumber lainnya, telah disebutkan dengan jelas di dalam daftar pustaka sesuai dengan kaidah penulisan ilmiah.

Saya menyadari bahwa dalam proses penyusunan Tugas Akhir ini, saya memanfaatkan bantuan teknologi kecerdasan buatan (AI) seperti *ChatGPT/Grammarly/Alat AI lain* hanya sebatas untuk:

1. Membantu dalam penyusunan tata bahasa, parafrasa, atau perbaikan redaksi.
2. Memberikan inspirasi ide awal atau kerangka berpikir, yang selanjutnya saya kembangkan secara mandiri.
3. Membantu dalam pengecekan konsistensi format, ejaan, dan tata tulis.

Saya menegaskan bahwa tidak ada bagian dari karya ilmiah ini yang seluruhnya dibuat oleh AI tanpa keterlibatan pemikiran kritis saya sendiri. Tanggung jawab penuh atas isi, keaslian, dan kebenaran karya ilmiah ini ada pada diri saya.

Apabila di kemudian hari terbukti bahwa pernyataan ini tidak benar, maka saya bersedia menerima segala konsekuensi sesuai dengan peraturan yang berlaku di Universitas Muhammadiyah Jember.

Jember, 31 Januari 2026
Yang membuat pernyataan,



Muhammad Ni'mad Rahmatullah

HALAMAN PENGESAHAN
PENERAPAN ALGORITMA MULTINOMIAL NAÏVE BAYES
UNTUK KLASIFIKASI SENTIMEN PADA HASHTAG
#KABURAJADULU DI MEDIA SOSIAL X

Oleh:

Muhammad Ni'mad Rahmatullah

2110651103

Telah mempertanggung jawabkan Laporan Tugas Akhirnya pada Sidang Tugas Akhir
31 Januari 2026 sebagai salah satu kelulusan dan mendapatkan gelar Sarjana

Komputer (S.Kom)

di

Universitas Muhammadiyah Jember

Disetujui oleh,

Dosen Penguji I



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

NIDN. 0714078704

Dosen Pembimbing I



Moh. Dasuki, M.Kom

NIDN. 0722109103

Dosen Penguji II



Dewi Lusiana Peter, MT.

NIDN. 0712086702

Dosen Pembimbing II



Habibatul Azizah Al Faruq, M.Pd

NIDN. 0718128901

Mengesahkan,
Dekan Fakultas Teknik



Prof. Dr. Ir. Muntar, S.T., M.T., IPM.
NIDN. 0010067301

Mengetahui,
Ketua Program Studi
Teknik Informatika



Rosita Yuniarti, S.Kom., M.Cs
NIDN. 0629018601

MOTTO

“The more you know, the more you realize you know nothing.”

(Socrates)

“Perceive that which cannot be seen with the eye.”

(Miyamoto Musashi)



KATA PENGANTAR

Puji Syukur penulis panjatkan kehadirat Allah S.W.T., atas segala rahmat dan hidayah-Nya yang telah dilimpahkan, sehingga penulis dapat menyelesaikan tugas akhir dengan judul **“PENERAPAN ALGORITMA MULTINOMIAL NAÏVE BAYES UNTUK KLASIFIKASI SENTIMEN PADA HASHTAG #KABURAJADULU DI MEDIA SOSIAL X”**. Penelitian tugas akhir ini mendasarkan pada isu ketenaga kerjaan di Indonesia. Tugas akhir ini merupakan karya ilmiah yang disusun dalam upaya untuk menyelesaikan pendidikan sarjana (S1) pada Fakultas Teknik Prodi Teknik Informatika Universitas Muhammadiyah Jember. Persembahkan tugas akhir ini dan rasa terimakasih saya ucapkan untuk:

1. Allah S.W.T., Yang Maha Pengasih dan Maha Penyayang, atas segala nikmat, rahmat, dan karunia-Nya sehingga penulis dapat menyelesaikan penelitian ini dan meraih gelar Sarjana Komputer (S.Kom) dengan lancar.
2. Ucapan terima kasih yang sebesar-besarnya penulis sampaikan kepada keluarga tercinta, Bapak Moch. Bisri dan Ibu Dra. Chomsatun, atas doa, dukungan, perhatian, dan kasih sayang yang senantiasa diberikan selama proses penyusunan karya ini.
3. Penulis sangat berterimakasih kepada Bapak Dasuki, M.Kom., selaku pembimbing utama, Ibu Habibatul Azizah Al Faruq, M.Pd selaku pembimbing kedua, atas segala perhatian dan bimbingannya serta arahan-arahan yang diberikan kepada penulis dalam upaya menyelesaikan tugas akhir ini.
4. Terimakasih penulis disampaikan pula kepada Bapak Ginanjar Abdurrahman, S.Si., M.Pd dan Ibu Dewi Lusiana Peter, MT. atas bantuan dan kesedian serta saran-saran yang diberikan kepada penulis dalam ujian tugas akhir.
5. Ucapan terimakasih yang sebesar-besarnya penulis sampaikan kepada Prof. Dr. Ir. Muhtar, S.T., M.T., IPM, selaku Dekan Fakultas Teknik, atas kesediaannya penulis belajar di Fakultas Teknik Prodi Teknik Informatika Universitas Muhammadiyah Jember.

6. Ucapan terima kasih juga penulis sampaikan kepada Kakak Moch. Khairul Anam atas bantuan finansial, doa, dan dukungan yang sangat berarti dalam penyelesaian skripsi ini.
7. Ucapan terima kasih juga penulis sampaikan kepada Ananda atas doa, bantuan dan dukungannya dalam penyelesaian skripsi ini.
8. Ucapan terima kasih juga kami sampaikan kepada seluruh teman-teman anggota grup *Gang-Gang* atas doa, bantuan, kebersamaan, dan dukungan yang diberikan dari awal perkuliahan hingga terselesaikannya skripsi ini.
9. Ucapan terima kasih juga penulis sampaikan kepada teman-teman KKN yang telah mendoakan serta memberikan bantuan dan dukungan dalam penyelesaian skripsi ini.
10. Ucapan terima kasih juga penulis sampaikan kepada seluruh pihak yang turut membantu dan memberikan dukungan, baik secara langsung maupun tidak langsung, dalam penyusunan skripsi ini.

Sebagai penutup, peneliti menyadari bahwa dalam penulisan skripsi ini masih jauh dari sempurna. Oleh karena itu, peneliti sangat mengharapkan kritik dan saran yang bersifat membangun dari berbagai pihak demi penyempurnaan karya ini di masa yang akan datang. Peneliti berharap semoga skripsi ini dapat memberikan manfaat serta menambah wawasan bagi pembaca dan pihak-pihak yang berkepentingan.

DAFTAR ISI

TUGAS AKHIR	i
PERNYATAAN ORISINALITAS	ii
HALAMAN PENGESAHAN	iii
MOTTO	iv
ABSTRAK	vi
KATA PENGANTAR	viii
DAFTAR ISI	ix
DAFTAR TABEL	xii
DAFTAR GAMBAR	xiii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Batasan Penelitian	3
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	4
BAB II TINJAUAN PUSTAKA	5
2.1 Penelitian terdahulu	5
2.2 Media Sosial sebagai Sumber Data <i>Text</i>	6
2.3 <i>Text Mining</i> dan Analisis Sentimen	7
2.4 <i>Missing Value</i>	8
2.5 <i>Preprocessing Text</i>	9
2.5.1 <i>Labeling</i>	9
2.5.2 <i>Cleaning data</i>	10
2.5.3 <i>Case Folding</i>	10
2.5.4 <i>Stopword Removal</i>	10
2.5.5 <i>Stemming</i>	11
2.5.6 <i>Normalization</i>	11

2.5.7 Tokenisasi	11
2.6 TF - IDF	12
2.7 Multinomial Naïve Bayes	12
2.8 Random Search	14
2.9 K-Fold Cross Validation.....	15
2.10 Validasi dan Evaluasi Model.....	15
BAB III METODE PENELITIAN.....	17
3.1 Desain Penelitian.....	17
3.2 Pengumpulan Data	18
3.3 Preprocessing Data	18
3.3.1 Labeling	18
3.3.2 Cleaning.....	19
3.3.3 Case Folding.....	19
3.3.4 Stopword Removal.....	20
3.3.5 Stemming.....	21
3.3.6 Normalization.....	21
3.3.7 Tokenisasi	22
3.4 TF-IDF	23
3.5 Implementasi Model.....	24
3.5.1 Smoothing.....	25
3.5.2 Multinomial Naïve Bayes	26
3.6 Model Training dan Validasi.....	28
3.6.1 Random Search	28
3.6.2 K-Fold Cross Validation.....	28
3.7 Evaluasi Model.....	29
BAB IV PEMBAHASAN DAN HASIL	32
4.1 Pengumpulan Data	32
4.2 Preprocessing.....	32
4.2.1 Labeling.....	32
4.2.2 Cleaning.....	33

4.2.3 Case Folding	34
4.2.4 Stopword Removal dan Stemming	36
4.2.5 Normalization	37
4.2.6 Tokenisasi	39
4.3 TF-IDF	42
4.4 Hasil Implementasi Metode	44
4.4.1 Smoothing	44
4.4.2 Multinomial Naive Bayes	45
4.5 Model Training dan Validasi	46
4.5.1 Random Search	47
4.5.2 K-Fold Cross Validation	48
4.6 Evaluasi Model	50
BAB V KESIMPULAN	54
5.1 Kesimpulan	54
5.2 Saran	54
DAFTAR PUSTAKA	55
LAMPIRAN-LAMPIRAN	62
CURRICULUM VITAE	63

DAFTAR TABEL

Tabel 2. 1 Penelitian Terdahulu.....	5
Tabel 3. 1 <i>Labeling</i>	18
Tabel 3. 2 <i>Data Cleaning</i>	19
Tabel 3. 3 <i>Case Folding</i>	20
Tabel 3. 4 <i>Stopword Removal</i>	20
Tabel 3. 5 <i>Stemming</i>	21
Tabel 3. 6 <i>Normalization</i>	22
Tabel 3. 7 <i>Tokenisasi</i>	22
Tabel 3. 8 <i>TF-IDF</i>	24
Tabel 3. 9 <i>Probabilitas Term ke-k Muncul di Kelas Positif</i>	27
Tabel 3. 10 <i>Probabilitas Term ke-k Muncul di Kelas Negatif</i>	27
Tabel 3. 11 <i>K-Fold Cross Validation</i>	29
Tabel 3. 12 <i>Data Contoh</i>	30
Tabel 4. 1 <i>Sampel Data</i>	50

DAFTAR GAMBAR

Gambar 2. 1 Ilustrasi <i>K-Fold Cross Validation</i>	15
Gambar 2. 2 Ilustrasi <i>Confusion Matrix</i>	16
Gambar 3. 1 Ilustrasi Alur Penelitian.....	17
Gambar 3. 2 <i>Confusion Matrix</i>	31
Gambar 4. 1 Code untuk menghilangkan Spam teks.....	32
Gambar 4. 2 Hasil <i>Labeling</i>	33
Gambar 4. 3 Hasil <i>Cleaning</i>	34
Gambar 4. 4 Hasil <i>Case Folding</i>	35
Gambar 4. 5 <i>Stopword Removal</i> dan <i>Stemming</i>	37
Gambar 4. 6 Hasil <i>Normalization</i>	38
Gambar 4. 7 Hasil <i>Tokenisasi</i>	40
Gambar 4. 8 Hasil <i>Word Cloud</i>	41
Gambar 4. 9 Code <i>Cleaning</i> secara manual.....	41
Gambar 4. 10 Hasil Data Kosong.....	42
Gambar 4. 11 Perhitungan TF-IDF.....	44
Gambar 4. 12 Hasil <i>Random Search</i>	48
Gambar 4. 13 Hasil <i>Random Search</i> dan <i>K-Fold Cross Validation</i>	50
Gambar 4. 14 Hasil <i>Confusion Matrix</i>	52
Gambar 4. 15 Hasil Evaluasi Model.....	53