

LAPORAN TUGAS AKHIR

PERBANDINGAN KINERJA ALGORITMA *NAÏVE BAYES* DAN *K-NN* DALAM ANALISIS SENTIMEN ULASAN APLIKASI *SIP* DISDUKCAPIL JEMBER

Diajukan Untuk Memenuhi Persyaratan Guna Meraih Sarjana Komputer
Teknik Informatika Universitas Muhammadiyah Jember



SANDI DWIKURNIAWAN

2210654012

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

2026

HALAMAN PERSETUJUAN SIDANG TUGAS AKHIR

PERBANDINGAN KINERJA ALGORITMA
NAÏVE BAYES DAN K-NN DALAM ANALISIS SENTIMEN
ULASAN APLIKASI SIP DISDUKCAPIL JEMBER

Oleh:

Sandi Dwikurniawan

2210654012

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

di

Universitas Muhammadiyah Jember

Disetujui Oleh,

**Pembimbing I**

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

Pembimbing II

Henny Wahyu Sulisty, S.Kom., M.Kom
NIDN. 0718128901

LEMBAR PENGESAHAN

**PERBANDINGAN KINERJA ALGORITMA
NAÏVE BAYES DAN K-NN DALAM ANALISIS SENTIMEN
ULASAN APLIKASI SIP DISDUKCAPIL JEMBER**

Oleh:

Sandi Dwikurniawan

2210654012

Telah mempertanggungjawabkan Laporan Tugas Akhirnya pada sidang Tugas Akhir Tanggal 24 Bulan Januari Tahun 2026 sebagai salah satu syarat kelulusan dan mendapat gelar Sarjana Komputer (S. Kom)

di

Universitas Muhammadiyah Jember

Disetujui Oleh,

**Dosen Penguji:
Penguji I**



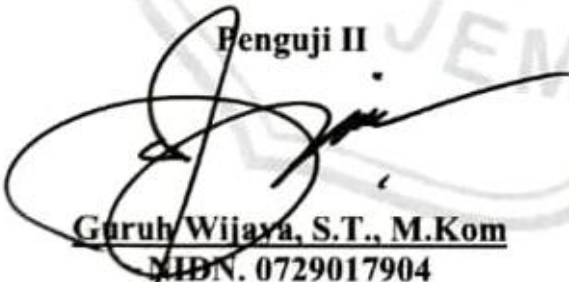
Nur Qodariyah Fitriyah, S.T., M.Kom
NIDN. 0705078006

**Dosen Pembimbing:
Pembimbing I**



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

Penguji II



Garuh Wijaya, S.T., M.Kom
NIDN. 0729017904

Pembimbing II



Henny Wahyu Sulistyo, S.Kom., M.Kom
NIDN. 0718128901

**Mengesahkan,
Dekan Fakultas Teknik**



Prof. Dr. Ir. Muhtar, ST., MT., IPM.
NIDN. 0010067301

**Mengetahui,
Ketua**

Program Studi Teknik Informatika



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

PERNYATAAN ORISINALITAS

Saya yang bertanda tangan di bawah ini:

Nama : Sandi Dwikurniawan
NIM : 2210654012
Program Studi : Teknik Informatika
Judul Tugas Akhir : Perbandingan Kinerja Algoritma *Naïve Bayes* dan *K-NN*
Dalam Analisis Sentimen Ulasan Aplikasi SIP
Disdukcipil Jember

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, 28 Januari 2026

Yang membuat pernyataan,



Sandi Dwikurniawan

MOTTO

“Belajar dari kemarin, hiduplah untuk hari ini, berharap untuk besok. Yang penting jangan berhenti bertanya. Curiosity punya alasan tersendiri untuk eksis”

– Albert Einstein

“Sains bukan hanya kumpulan pengetahuan, melainkan cara berpikir.”

– Carl Sagan

“Tanpa data, seseorang hanyalah memiliki pendapat.”

– W. Edwards Deming



PERSEMBAHAN

Puji dan syukur penulis panjatkan ke hadirat Allah SWT atas limpahan rahmat, hidayah, dan karunia-Nya sehingga penulis dapat menyelesaikan karya ilmiah ini dengan baik.

Tugas Akhir ini penulis persembahkan dengan penuh rasa hormat dan terima kasih yang setulus-tulusnya kepada:

1. Ayah tercinta (Almarhum), yang semasa hidupnya telah menjadi teladan, sumber kekuatan, dan inspirasi bagi penulis. Doa dan perjuangan beliau akan selalu hidup dalam setiap langkah yang penulis tempuh. Semoga Allah SWT menempatkan beliau di tempat terbaik di sisi-Nya.
2. Ibu tercinta, yang senantiasa memberikan doa yang tiada henti, kasih sayang, dukungan moral maupun material, serta menjadi pilar utama dalam perjalanan pendidikan dan kehidupan penulis.
3. Keluarga besar penulis, yang selalu memberikan perhatian, semangat, dan dorongan untuk terus berusaha dan tidak mudah menyerah.
4. Dosen pembimbing dan dosen penguji, atas bimbingan, arahan, serta masukan yang sangat berarti selama proses penyusunan karya ilmiah ini.
5. Seluruh dosen dan civitas akademika, yang telah memberikan ilmu pengetahuan dan pengalaman berharga selama masa perkuliahan.
6. Teman-teman seperjuangan, yang telah menemani, membantu, dan memberikan semangat dalam suka maupun duka selama proses studi.
7. Dinas Kependudukan dan Pencatatan Sipil Kabupaten Jember, atas ketersediaan dan bantuannya dalam pemenuhan data pendukung pada penelitian ini.
8. Semua pihak yang tidak dapat disebutkan satu per satu, yang telah memberikan dukungan secara langsung maupun tidak langsung.

Semoga karya ilmiah ini dapat memberikan manfaat dan menjadi bentuk pengabdian penulis kepada orang tua, keluarga, serta masyarakat.

KATA PENGANTAR

Puji Syukur penulis panjatkan kehadiran Allah S.W.T., atas segala rahmat dan hidayah-Nya yang telah dilimpahkan, sehingga penulis dapat menyelesaikan tugas akhir dengan judul: Perbandingan Kinerja Algoritma *Naïve Bayes* dan *K-NN* Dalam Analisis Sentimen Ulasan Aplikasi SIP Disdukcapil Jember.

Penelitian tugas akhir ini mendasarkan pada isu peningkatan kualitas pelayanan publik, khususnya dalam penyelenggaraan administrasi kependudukan yang memerlukan evaluasi berkelanjutan berdasarkan umpan balik masyarakat. Tugas akhir ini merupakan karya ilmiah yang disusun dalam upaya untuk menyelesaikan pendidikan sarjana (S1) pada Fakultas Teknik Prodi Teknik Informatika Universitas Muhammadiyah Jember.

Penulis sangat berterimakasih kepada Ibu Rosita Yanuarti, S.Kom., M.Cs. selaku pembimbing utama, Bapak Henny Wahyu Sulisty, S.Kom., M.Kom, selaku pembimbing kedua, atas segala perhatian dan bimbingannya serta arahan-arahan yang diberikan kepada penulis dalam upaya menyelesaikan tugas akhir ini.

Terimakasih penulis disampaikan pula kepada Ibu Nur Qodariyah Fitriyah, S.T., M.Kom dan Bapak Guruh Wijaya, S.T., M.Kom, atas bantuan dan kesedian serta saran-saran yang diberikan kepada penulis dalam ujian tugas akhir.

Ucapan terimakasih yang sebesar-besarnya penulis sampaikan kepada Prof. Dr. Ir. Muhtar, ST., MT., IPM. selaku Dekan Fakultas Teknik, atas kesediaanya penulis belajar di Fakultas Teknik Prodi Teknik Informatika Universitas Muhammadiyah Jember.

Tidak lupa mengucapkan banyak terimakasih kepada Disdukcapil Kabupaten Jember yang telah memberikan bantuan data dan informasi selama pelaksanaan penelitian lapangan.

Harapan penulis semoga laporan hasil penelitian tugas akhir ini bisa bermanfaat bagi pembaca dan berguna bagi pengembangan Ilmu Informatika.

DAFTAR ISI

HALAMAN PERSETUJUAN SIDANG TUGAS AKHIR	i
LEMBAR PENGESAHAN	ii
PERNYATAAN ORISINALITAS	iii
MOTTO	iv
PERSEMBAHAN	v
ABSTRAK	vi
<i>ABSTRACT</i>	vii
KATA PENGANTAR.....	viii
DAFTAR ISI.....	ix
DAFTAR GAMBAR.....	xii
DAFTAR TABEL	xiii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	3
1.3 Tujuan Penelitian.....	3
1.4 Manfaat Penelitian	3
1.5 Batasan Penelitian.....	4
BAB II TINJAUAN PUSTAKA	5
2.1 Dinas Kependudukan dan Pencatatan Sipil	5
2.2 <i>SIP</i> Disdukcapil Kabupaten Jember	6
2.3 Analisis Sentimen	6
2.4 <i>Naïve Bayes</i>	7
2.5 <i>K-Nearest Neighbors (K-NN)</i>	8
2.6 <i>Web Scraping</i>	9
2.7 <i>Preprocessing Text</i>	9

2.8	<i>Vectorizer</i>	10
2.9	<i>ADASYN</i>	10
2.10	<i>Confusion Matrix</i>	12
2.11	<i>Recall</i>	13
2.12	<i>Precision</i>	13
2.13	<i>F2-Score</i>	13
2.14	Penelitian Terdahulu.....	14
BAB III METODOLOGI PENELITIAN		17
3.1	<i>Scraping</i>	17
3.2	<i>Data Understanding</i>	19
3.3	<i>Preprocessing Text</i>	20
3.3.1	<i>Data Cleaning</i>	20
3.3.2	<i>Lowercasing</i>	23
3.3.3	<i>Word Normalizer</i>	24
3.3.4	<i>Tokenizing</i>	25
3.3.5	<i>Stop Words Removal</i>	26
3.3.6	<i>Stemming</i>	27
3.4	<i>Data preparation</i>	28
3.4.1	<i>Pelabelan Data</i>	29
3.4.2	<i>Splitting Data</i>	30
3.4.3	<i>Vectorizer</i>	30
3.5	<i>ADASYN</i>	33
3.6	<i>Modelling</i>	34
3.6.1	<i>Naïve Bayes</i>	35
3.6.2	<i>K-NN</i>	38
3.6.3	<i>Implementasi Naïve Bayes dan K-NN menggunakan Python</i>	41
3.7	<i>Analisa Perbandingan Algoritma</i>	42
3.7.1	<i>Konsep Analisa Confusion Matrix</i>	42
3.7.2	<i>Konsep Analisa Hasil Klasifikasi</i>	43
3.7.3	<i>Konsep Analisa F2-Score</i>	43

BAB IV HASIL DAN PEMBAHASAN	44
4.1 Hasil <i>Scraping</i> Data	44
4.2 Analisis <i>Data Understanding</i>	44
4.3 Hasil Proses <i>Preprocessing Text</i>	45
4.3.1 Hasil Proses <i>Data Cleaning</i>	46
4.3.2 Hasil Proses <i>Lowercasing</i>	47
4.3.3 Hasil Proses <i>Word Normalizer</i>	47
4.3.4 Hasil Proses <i>Tokenizing</i>	48
4.3.5 Hasil Proses <i>Stop Word Removal</i>	48
4.3.6 Hasil Proses <i>Stemming</i>	49
4.4 Hasil Proses <i>Data Preparation</i>	50
4.4.1 Hasil Proses Pelabelan Data	51
4.4.2 Hasil Proses <i>Splitting Data</i>	51
4.4.3 Hasil Proses <i>Vectorizer</i>	52
4.5 Hasil Proses <i>ADASYN</i>	52
4.6 Hasil Pemodelan Algoritma Klasifikasi	53
4.6.1 Hasil Pemodelan <i>Naïve Bayes</i>	53
4.6.2 Hasil Pemodelan <i>K-NN</i>	54
4.7 Perbandingan Model <i>Naïve Bayes</i> dengan <i>K-NN</i>	56
BAB V KESIMPULAN DAN SARAN.....	58
5.1 Kesimpulan	58
5.2 Saran	58
DAFTAR PUSTAKA.....	59
LAMPIRAN.....	62

DAFTAR GAMBAR

Gambar 3.1 Flowchart Metodologi Penelitian	17
Gambar 3.2 Tampilan <i>page</i> IKM <i>website</i> SIP	18
Gambar 3.3 Struktur halaman <i>website</i> SIP	19
Gambar 3.4 <i>Preprocessing Text</i>	20
Gambar 3.5 Ketidakseimbangan Data.....	33
Gambar 3.6 <i>Flowchart</i> proses pemodelan klasifikasi	41
Gambar 3.7 Ilustrasi <i>confusion matrix</i>	42
Gambar 4.1 struktur <i>HTML website</i> SIP	44
Gambar 4.2 Contoh <i>dataset</i> masih kotor.....	46
Gambar 4.3 Contoh hasil <i>data cleaning</i>	47
Gambar 4.4 Contoh hasil <i>lowercasing</i>	47
Gambar 4.5 Contoh hasil proses tokenisasi	48
Gambar 4.6 Contoh hasil proses <i>stop word removal</i>	49
Gambar 4.7 Contoh hasil proses <i>stemming</i>	50
Gambar 4.8 Contoh hasil proses pengembalian data ke bentuk normal	50
Gambar 4.9 Hasil proses pelabelan.....	51
Gambar 4.10 Contoh hasil proses <i>vetrozier</i>	52
Gambar 4.11 Hasil proses <i>ADASYN</i>	52
Gambar 4. 12 Contoh Perbandingan Hasil Klasifikasi	57

DAFTAR TABEL

Tabel 2.1 Daftar penelitian terdahulu.....	15
Tabel 3.1 Contoh Data Sampel Data Mentah.....	21
Tabel 3.2 Hasil <i>Data Cleaning</i>	22
Tabel 3.3 Contoh hasil penerapan <i>lowercasing</i>	23
Tabel 3.4 Contoh hasil penerapan <i>word normalizer</i>	25
Tabel 3.5 Contoh hasil penerapan <i>tokenizing</i>	25
Tabel 3.6 Contoh hasil penerapan <i>stop word removal</i>	27
Tabel 3.7 Contoh hasil penerapan <i>stemming</i>	27
Tabel 3.8 Contoh data bersih hasil <i>preprocessing text</i>	29
Tabel 3.9 Contoh penerapan pelabelan	29
Tabel 3.10 Contoh Sampel Data Vektorisasi Tokenisasi.....	30
Tabel 3.11 Contoh Perhitungan <i>TF</i>	31
Tabel 3.12 Contoh Perhitungan <i>IDF</i>	31
Tabel 3.13 Contoh Hasil Perhitungan <i>TF-IDF</i>	32
Tabel 3. 14 Hasil Klasifikasi tanpa <i>ADASYN</i>	34
Tabel 3.15 Hasil Perhitungan <i>TF-IDF</i> sentimen positif.....	35
Tabel 3.16 Hasil Perhitungan <i>TF-IDF</i> sentimen negatif.....	36
Tabel 3.17 perhitungan <i>TF-IDF</i> pada dokumen “cukup cepat”	38
Tabel 3.18 Perangkingan jarak <i>euclidean</i>	40
Tabel 3.20 Tabel Perbandingan Kesalahan <i>Confusion Matrix</i>	42
Tabel 3.19 Konsep tabel perbandingan hasil klasifikasi.....	43
Tabel 3.21 Interpretasi nilai <i>F2-Score</i>	43
Tabel 4.1 Identifikasi permasalahan dalam <i>dataset</i>	45
Tabel 4.2 Contoh normalisasi kata.....	48
Tabel 4.4 <i>Confusion Matrix Naïve Bayes</i>	53
Tabel 4.3 <i>Classification Reports Naïve Bayes</i>	53
Tabel 4.5 Hasil pemodelan <i>K-NN</i> Skenario A	54
Tabel 4.6 Hasil pemodelan <i>K-NN</i> Skenario B	55
Tabel 4.7 Hasil pemodelan <i>K-NN</i> Skenario C	55
Tabel 4.8 Hasil pemodelan <i>K-NN</i> Skenario D	55

Tabel 4.9 <i>Confusion Matrix K-NN</i>	56
Tabel 4.10 Perbandingan Hasil Kedua Model Klasifikasi	57

