

**ANALISIS SENTIMEN KOMENTAR PENGGUNA DI MEDIA SOSIAL
TENTANG DINASTI POLITIK DENGAN PENDEKATAN NAIVE BAYES
DAN K-NN**

SKRIPSI SARJANA SISTEM INFORMASI



Oleh:

Tio Adi Kurniawan

227006426140

FAKULTAS TEKNOLOGI KOMUNIKASI DAN INFORMATIKA

PROGRAM STUDI SISTEM INFORMASI
UNIVERSITAS NASIONAL 2025

HALAMAN PENGESAHAN
TUGAS AKHIR

ANALISIS SENTIMEN KOMENTAR PENGGUNA DI MEDIA SOSIAL
TENTANG DINASTI POLITIK DENGAN PENDEKATAN NAIVE BAYES
DAN K-NN



Dosen Pembimbing 1

(Dr. Aris Gunaryati, S.SI., MMSI)

Dosen Pembimbing 2

()

PERNYATAAN KEASLIAN TUGAS AKHIR

Saya menyatakan dengan sesungguhnya bahwa Tugas Akhir dengan judul :

ANALISIS SENTIMEN KOMENTAR PENGGUNA DI MEDIA SOSIAL TENTANG DINASTI POLITIK DENGAN PENDEKATAN NAIVE BAYES DAN K-NN

Yang dibuat untuk melengkapi salah satu persyaratan menjadi Sarjana Komputer pada Program Studi Sistem Informasi Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional, sebagaimana yang saya ketahui adalah bukan merupakan tiruan atau publikasi dari Tugas Akhir yang pernah diajukan atau dipakai untuk mendapatkan gelar di lingkungan Universitas Nasional maupun perguruan tinggi atau instansi lainnya, kecuali pada bagian – bagian tertentu yang menjadi sumber informasi atau acuan yang dicantumkan sebagaimana mestinya.

Jakarta, 28 Februari 2025



Tio Adi Kurniawan
227006426140

LEMBAR PERSETUJUAN REVIEW AKHIR

Tugas Akhir dengan judul :

ANALISIS SENTIMEN KOMENTAR PENGGUNA DI MEDIA SOSIAL TENTANG DINASTI POLITIK DENGAN PENDEKATAN NAIVE BAYES DAN K-NN

Dibuat untuk melengkapi salah satu persyaratan menjadi Sarjana Komputer pada Program Studi Sistem Informasi, Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional. Tugas Akhir ini diujikan pada Sidang Review Akhir Semester Ganjil 2024-20225 pada tanggal 26 Februari Tahun 2025

Dosen Pembimbing 1



Dr. Aris Gunaryati, S.SI., MMSI
0313087705

Ketua Program Studi



Dr. Andrianingsih, S.Kom., MMSI
0303097902

LEMBAR PERSETUJUAN JUDUL YANG TIDAK ATAU YANG DIREVISI

Nama : Tio Adi Kurniawan

NPM : 227006426140

Fakultas/Akademi : Fakultas Teknologi Komunikasi dan Informatika

Program Studi : Sistem Informasi

Tanggal Sidang : 26 Februari 2025

JUDUL DALAM BAHASA INDONESIA :

ANALISIS SENTIMEN KOMENTAR PENGGUNA DI MEDIA SOSIAL TENTANG
DINASTI POLITIK DENGAN PENDEKATAN NAIVE BAYES DAN K-NN

JUDUL DALAM BAHASA INGGRIS :

SENTIMENT ANALYSIS USER COMMENTS ON SOCIAL MEDIA ABOUT THE
POLITICAL DYNASTY USING THE NAIVE BAYES AND K-NN

TANDA TANGAN DAN TANGGAL

Pembimbing 1	Ka. Prodi	Mahasiswa
TGL : 28 Februari 2025	TGL : 28 Februari 2025	TGL :
 Dr. Aris Gunartati, S.Si, MMSi	 Dr. Andrianingsih, S.Pd., M.Pd.	

KATA PENGANTAR

Puji syukur ke hadirat Tuhan Yang Maha Esa atas segala rahmat dan karuniaNya sehingga skripsi ini dapat diselesaikan dengan baik. Skripsi yang berjudul "

ANALISIS SENTIMEN KOMENTAR PENGGUNA DI MEDIA SOSIAL TENTANG DINASTI POLITIK DENGAN PENDEKATAN NAIVE BAYES DAN K-NN" disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana di Fakultas Teknologi Komunikasi dan Informatika, Program Studi Sistem Informasi, Universitas Nasional.

Penyusunan skripsi ini tidak terlepas dari bantuan, bimbingan, serta dukungan dari berbagai pihak. Oleh karena itu, dengan segala hormat dan rasa terima kasih yang mendalam, penulis menyampaikan apresiasi kepada:

1. Bapak Dr. El Amry Bermawi Putera, M.A., selaku Rektor Universitas Nasional yang telah memberikan fasilitas dan kesempatan bagi mahasiswa dalam menempuh pendidikan tinggi dengan baik.
2. Bapak Dr. Agung Triayudi, S.Kom., M.Kom, selaku Dekan Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional yang telah memberikan kesempatan dan fasilitas dalam penyelesaian studi.
3. Ibu Dr. Andrianingsih, S.Kom.,MMSI., selaku Ketua Program Studi Sistem Informasi Universitas Nasional atas arahan dan kebijakan yang diberikan selama masa perkuliahan.
4. Ibu Dr Aris Gunaryati, S.Kom.,MMSI., selaku dosen pembimbing yang telah memberikan bimbingan, masukan, serta motivasi yang sangat berharga dalam proses penyusunan skripsi ini.
5. Bapak/Ibu Dosen Universitas Nasional, yang telah memberikan ilmu, wawasan, dan pengalaman berharga selama masa perkuliahan.
6. Keluarga tercinta ibu, ayah, kakak, adik, atas segala doa, dukungan, serta pengorbanan yang tak ternilai harganya dalam mendukung penyelesaian studi ini.

7. Teman-teman seperjuangan Isnando, Fauzan, Irzan, Aji, Rendi, Putra, Rival, Lintar yang selalu memberikan semangat, motivasi, serta dukungan moral dalam menjalani proses akademik dan penelitian ini.
8. Pasangan tercinta Shyfa Afrely Darma yang selalu hadir memberikan semangat, kesabaran, dukungan emosional, serta kepercayaan yang besar dalam perjalanan akademik ini. Terima kasih atas segala doa, motivasi, dan pengorbanan yang diberikan selama proses penyusunan skripsi ini.
9. Seluruh pihak yang tidak dapat disebutkan satu per satu, yang telah membantu secara langsung maupun tidak langsung dalam penyelesaian skripsi ini.

Penulis menyadari bahwa karya ini dapat memberikan manfaat bagi perkembangan ilmu pengetahuan, khususnya dalam bidang analisis sentimen dan pemanfaatan algoritma pembelajaran mesin dalam dunia digital. Skripsi ini masih memiliki keterbatasan dan kekurangan. Oleh karena itu, segala kritik dan saran yang membangun sangat diharapkan guna penyempurnaan penelitian ini.

Jakarta, 26 Februari 2025

Peneliti

Tio Adi Kurniawan

Abstrak

Fenomena dinasti politik di Indonesia telah menjadi isu yang kompleks dan sering kali menimbulkan kejadian di ruang publik, terutama di media sosial seperti YouTube dan TikTok. Perkembangan informasi teknologi telah memungkinkan masyarakat untuk lebih aktif dalam menyampaikan pendapat mereka terkait isu-isu politik, termasuk keinginan berbasis kekerabatan. Meskipun dinasti politik memiliki dasar sejarah yang kuat, praktik ini sering kali dikritik karena dianggap menghambat demokrasi dan meritokrasi. Oleh karena itu, analisis terhadap opini publik mengenai dinasti politik menjadi sangat relevan untuk memahami berbagai macam hal

Penelitian ini bertujuan untuk menganalisis sentimen komentar pengguna media sosial terhadap dinasti politik dengan membandingkan dua algoritma *machine learning*, yaitu *Naïve Bayes* dan *K-Tetangga Terdekat (K-NN)*. Guru Gembul dan praproses (pembelajaran mesin).

Model evaluasi dilakukan berdasarkan pengukuran akurasi, presisi, recall, dan *F1-score*. Untuk *Naïve Bayes* memiliki skor *F1* sebesar

Perbedaan distribusi sentimen antara kedua platform dapat diinterpretasikan dari karakteristik pengguna dan pola komunikasi yang berbeda. YouTube sebagai platform berbasis video panjang cenderung menarik diskusi yang lebih mendalam dan kritis, sementara TikTok yang berorientasi pada konten singkat cenderung memiliki komentar yang lebih singkat dan spontan. Selain itu, hasil penelitian ini memberikan wawasan mengenai efektivitas pembelajaran mesin algoritma dalam analisis sentimen berbasis

Kata Kunci : *Analisis Sentimen, Dinasti Politik, Media Sosial, *Naïve Bayes*, *Analisis Sentimen, Dinasti Politik, Media Sosial, Naïve Bayes, K-Nearest Neighbors (K-NN), Machine Learning*



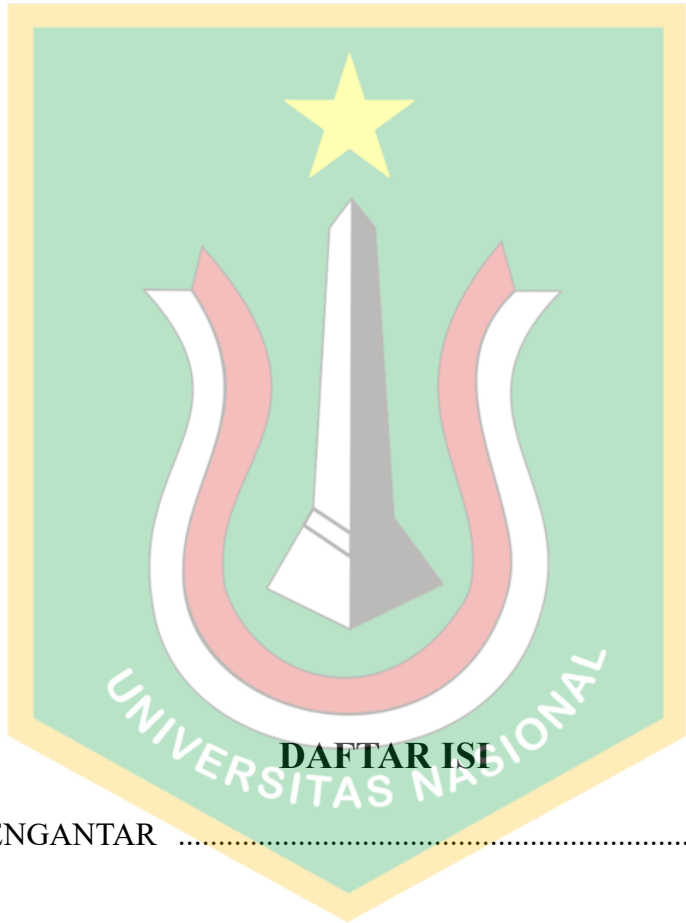
The phenomenon of political dynasties in Indonesia has become a complex issue that often sparks debates in public discourse, particularly on social media platforms such as YouTube and TikTok. This study aims to analyze user sentiment on social media regarding political dynasties by comparing two machine learning algorithms: *Naïve Bayes* and *K-Nearest Neighbors (K-NN)*. Guru Gembul Y

The models were evaluated using accuracy, precision, recall, and *F1-score* metrics. The *Naïve Bayes* algorithm *F1-score* of *F1-score* of 0

The differences in sentiment distribution across these platforms can be attributed to variations in user demographics and communication styles. YouTube, as a long-form video platform, tends to encourage more in-depth and critical discussions, whereas

TikTok, with its short-form content orientation, generates more concise and spontaneous comments. Furthermore, this study provides insights into the effectiveness of machine learning models in sentiment analysis and their applicability in various domains, particularly in understanding public opinion o

Keywords: *Sentiment Analysis, Political Dynasty, Social Media, Naïve Bayes, KNearest Neighbors (K-NN), Machine Learning*



DAFTAR ISI

KATA PENGANTAR	vi
Abstrak	ii
Abstract	ix
BAB I Pendahuluan	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan Penelitian	2
1.4 Batasan Masalah	3

1.5 Manfaat Penelitian	4
1.6 Sistematika Penulisan	5
BAB II	6
TINJAUAN PUSTAKA	6
2.1 Studi Literatur	6
2.2 Landasan Teori	9
2.2.1 Analisis Sentimen	9
2.2.2 Dinasti Politik	10
2.2.3 Media sosial yang di gunakan	11
2.2.4 Teknik Pengumpulan Data	11
2.2.4 Naive Bayes Classification	12
2.2.5 K-Nearest Neighbour	13
2.2.6 Machine Learning	14
2.2.7 Google Colab	14
BAB III METODOLOGI PENELITIAN	15
3.1 Lokasi Penelitian	15
3.2 Alur Penelitian	16
3.3 Pengumpulan Data	16
3.4 Preprocessing	18
3.5 Pelabelan data	19
3.5.1 Kamus Sentimen	19
3.5.2 Logika Pelabelan	20
3.6 Ekstraksi Fitur	20
3.7 Metode Klasifikasi	21

3.7.1 Naive Bayes	21
3.7.2 K-Nearest Neighbors (KNN)	21
3.8 Evaluasi Model	21
3.8.1 Tujuan Evaluasi Model	21
3.8.2 Metrik Evaluasi	22
3.8.3 Pengukuran Evaluasi	23
3.8.4 Cross-Validation	23
3.8.5 Proses Evaluasi	24
3.8.6 Perbandingan Model	24
3.8.7 Confusion Matrix	24
BAB IV HASIL DAN PEMBAHASAN	26
4.1 Pendahuluan	26
4.2 Proses Penelitian	26
4.2.1 Pengambilan Data	26
4.2.2 Preprocessing	28
4.2.3 Pelabelan Data	31
4.2.4 Distribusi Sentimen Youtube dan TikTok	32
4.2.5 Confusion Matrix	34
4.2.6 Perhitungan Manual	37
4.5 Perbandingan Performa Model	40
4.6 Analisis Kesalahan Klasifikasi	42
BAB V KESIMPULAN DAN SARAN	43
5.1 Kesimpulan	43
5.2 Saran	45
Daftar Pustaka	47
LAMPIRAN	
49 DAFTAR GAMBAR	

Gambar 3. 1 Alur Penelitian	16
Gambar 4. 1 Code Pengambilan Data	26
Gambar 4. 2 Tampilan Data	27
Gambar 4. 3 Pengambilan data TikTok	27
Gambar 4. 4 Clean Text	28
Gambar 4. 5 Normalisasi	29
Gambar 4. 6 Stopwords	30
Gambar 4. 7 Stemming	30
Gambar 4. 8 Pelabelan Data	32
Gambar 4. 9 Distribusi Sentimen Youtube	32
Gambar 4. 10 Distribusi Sentimen TikTok	33
DAFTAR TABEL	
Tabel 2. 1 Studi Literatur	6
Tabel 3. 1 Contoh Struktur Data	17
Tabel 3. 2 Preprocessing	18
Tabel 4. 1 Confusion Matrix Naive Bayes (Youtube)	37
Tabel 4. 2 Confusion Matrix KNN (Youtube)	38
Tabel 4. 3 Confusion Matrix Naive Bayes (TikTok)	38
Tabel 4. 4 Confusion Matrix KNN (TikTok)	39
Tabel 4. 5 Perbandingan Model	40
Tabel 4. 6 Kesalahan Klasifikasi Langkah	42