

**DETEKSI KOMENTAR CYBERBULLYING PADA TIKTOK
MENGUNAKAN ALGORITMA K-NEAREST NEIGHBOR (KNN)**

TUGAS AKHIR

Diajukan sebagai salah satu syarat memperoleh gelar Sarjana

Program Studi S-1 Teknologi Informasi



Disusun oleh:

NANDA SAPUTRI WINDI OKTAVIA

212104007

**PROGRAM STUDI S-1 TEKNOLOGI INFORMASI
FAKULTAS TEKNIK & TEKNOLOGI INFORMASI
UNIVERSITAS JENDERAL ACHMAD YANI YOGYAKARTA**

2025

HALAMAN PENGESAHAN

TUGAS AKHIR
DETEKSI KOMENTAR *CYBERBULLYING* PADA TIKTOK
MENGGUNAKAN ALGORITMA *K-NEAREST NEIGHBOR* (KNN)

Diajukan oleh:

NANDA SAPUTRI WINDI OKTAVIA

212104007

Telah dipertahankan di depan dewan penguji dan dinyatakan sah
sebagai salah satu syarat untuk memperoleh gelar Sarjana
di Fakultas Teknik & Teknologi Informasi
Universitas Jenderal Achmad Yani Yogyakarta

Tanggal: 09 Juli 2025

Mengesahkan:

Pembimbing

Alfina Rizqi Lahitani, S.Kom., M.Eng.

NIDN: 0506019202

Penguji

Rama Sahtyawan, S.T., M.Cs.

NIDN: 0518108001

Ketua Program Studi S-1 Teknologi Informasi
Fakultas Teknik & Teknologi Informasi
Universitas Jenderal Achmad Yani Yogyakarta



Rama Sahtyawan, S.T., M.Cs.

NPP: 2019.13.00150

PERNYATAAN

Saya yang bertanda tangan di bawah ini, adalah mahasiswa Fakultas Teknik dan Teknologi Informasi Universitas Jenderal Achmad Yani Yogyakarta,

Nama : Nanda Saputri Windi Oktavia
NPM : 212104007
Program Studi : S-1 Teknologi Informasi
Judul Tugas Akhir : Deteksi Komentar *Cyberbullying* pada TikTok
Menggunakan Algoritma *K-Nearest Neighbor* (KNN)

Menyatakan bahwa hasil penelitian dengan judul tersebut di atas adalah asli karya saya sendiri dan bukan hasil plagiarisme. Semua referensi dan sumber terkait yang dikutip dalam karya ilmiah ini telah ditulis sesuai kaidah penulisan ilmiah yang berlaku. Dengan ini, saya menyatakan untuk menyerahkan hak cipta penelitian kepada Universitas Jenderal Achmad Yani Yogyakarta guna kepentingan ilmiah.

Demikian surat pernyataan ini dibuat dengan sebenar-benarnya tanpa ada paksaan dari pihak mana pun. Apabila terdapat kekeliruan atau ditemukan adanya pelanggaran akademik di kemudian hari, maka saya bersedia menerima konsekuensi yang berlaku sesuai ketentuan akademik.

Yogyakarta, 03 Juli 2025



Nanda Saputri Windi Oktavia

KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Allah SWT atas limpahan rahmat, karunia, dan kemudahan-Nya sehingga penulis dapat menyelesaikan laporan skripsi dengan judul “Deteksi Komentar *Cyberbullying* Pada TikTok Menggunakan Algoritma *K-Nearest Neighbor* (KNN)”. Penyusunan skripsi ini merupakan salah satu syarat untuk menyelesaikan studi di Program Studi S-1 Teknologi Informasi, Fakultas Teknik dan Teknologi Informasi, Universitas Jenderal Achmad Yani Yogyakarta. Laporan ini dapat diselesaikan atas bimbingan, arahan, dan bantuan dari berbagai pihak. Pada kesempatan ini penulis dengan rendah hati mengucapkan terima kasih dengan setulus-tulusnya kepada:

1. Bapak Aris Wahyu Murdiyanto, S.Kom., M.Cs. selaku Dekan Fakultas Teknik dan Teknologi Informasi Universitas Jenderal Achmad Yani Yogyakarta;
2. Bapak Rama Sahtyawan, S.T., M.Cs.. selaku Ketua Program Studi S-1 Teknologi Informasi Fakultas Teknik dan Teknologi Informasi Universitas Jenderal Achmad Yani Yogyakarta;
3. Ibu Alfirma Rizqi Lahitani, S.Kom., M.Eng. selaku Dosen Pembimbing Tugas Akhir yang telah mendukung dan membantu selama berjalannya penelitian hingga penyusunan laporan;
4. Kedua orang tua, sosok luar biasa yang senantiasa menjadi sumber kekuatan dan penyemangat dalam menghadapi berbagai tantangan kehidupan. Dengan penuh cinta dan ketulusan, mereka tidak pernah lelah memberikan doa, kasih sayang, serta motivasi yang tulus dalam setiap langkah yang penulis tempuh. Terima kasih atas segala bentuk pengorbanan, kesabaran, dan kepercayaan yang telah diberikan kepada penulis selama proses studi ini berlangsung. Nasihat, dukungan, dan keyakinan yang selalu mereka tanamkan menjadi fondasi penting bagi penulis untuk terus maju hingga akhirnya dapat menyelesaikan pendidikan ini. Semoga kesehatan, kebahagiaan, dan umur panjang senantiasa menyertai keduanya, agar tetap

menjadi bagian dari setiap proses dan pencapaian dalam kehidupan penulis di masa yang akan datang;

5. Para dosen Fakultas Teknologi Informasi Universitas Jenderal Achmad Yani Yogyakarta, yang telah membagikan ilmu, wawasan, serta pengalaman berharga selama penulis menempuh studi;
6. Keluarga dan sahabat-sahabat penulis yang senantiasa memberikan semangat, motivasi, dan doa sehingga laporan tugas akhir ini dapat diselesaikan dengan lancar dan tepat waktu;
7. Rekan-rekan mahasiswa Program Studi S-1 Teknologi Informasi angkatan 2021 yang selalu memberi semangat dan dukungan selama menempuh pendidikan di Universitas Jenderal Achmad Yani Yogyakarta;
8. Semua pihak yang tidak dapat penulis sebutkan satu per satu, yang telah memberikan dukungan, serta dorongan baik secara langsung maupun tidak langsung selama proses penyusunan tugas akhir ini. Ucapan terima kasih juga kepada diri saya sendiri, atas komitmen, ketekunan, serta upaya yang terus dijaga dalam menghadapi setiap proses dan tantangan hingga tugas akhir ini dapat diselesaikan dengan baik.

Penulis menyadari bahwa penyusunan laporan tugas akhir ini masih memiliki berbagai kekurangan, mengingat keterbatasan pengetahuan, pengalaman, dan wawasan yang dimiliki. Oleh karena itu, dengan penuh kerendahan hati, penulis sangat menghargai setiap bentuk kritik dan saran yang bersifat membangun dari para pembaca demi penyempurnaan di masa yang akan datang. Penulis berharap laporan ini dapat memberikan manfaat, baik sebagai bahan referensi maupun sebagai tambahan wawasan, khususnya bagi penulis sendiri dan secara umum bagi siapa pun yang tertarik pada topik ini.

Yogyakarta, 03 Juli 2025

Nanda Saputri Windi Oktavia

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PENGESAHAN	ii
PERNYATAAN.....	iii
KATA PENGANTAR.....	iv
DAFTAR ISI.....	vi
DAFTAR TABEL	viii
DAFTAR GAMBAR.....	ix
DAFTAR LAMPIRAN	x
DAFTAR SINGKATAN	xi
INTISARI	xii
ABSTRACT	xiii
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah	4
1.3 Pertanyaan Penelitian	4
1.4 Tujuan Penelitian	4
1.5 Manfaat Hasil Penelitian	4
BAB 2 TINJAUAN PUSTAKA DAN LANDASAN TEORI	6
2.1 Tinjauan Pustaka	6
2.2 Landasan Teori.....	13
2.3 <i>Cyberbullying</i> pada Media Sosial	14
2.3.1 <i>Cyberbullying</i>	14
2.3.2 Media Sosial	14
2.4 <i>Machine Learning</i>	14
2.4.1 Pengertian <i>Machine Learning</i>	14
2.4.2 <i>Supervised Learning</i>	15
2.4.3 Klasifikasi.....	15
2.5 Pendeteksian <i>Cyberbullying</i>	15
2.6 Algoritma <i>K-Nearest Neighbor</i> (KNN)	16

2.7 Kinerja Algoritma	16
2.7.1 <i>Confusion Matrix</i>	16
BAB 3 METODE PENELITIAN	18
3.1 Bahan dan Alat Penelitian	25
3.2 Jalan Penelitian.....	26
BAB 4 HASIL PENELITIAN	28
4.1 Ringkasan Hasil Penelitian	28
4.2 Pembahasan.....	28
4.2.1 Pengumpulan Data	28
4.2.2 Pembersihan Data Awal	31
4.2.3 <i>Pre-Processing</i> Data	32
4.2.4 Pelabelan Data.....	41
4.2.5 Pembobotan TF-IDF	45
4.2.6 <i>Splitting</i> Data.....	48
4.2.7 Klasifikasi Model <i>K-Nearest Neighbor</i> (KNN)	49
4.2.8 Evaluasi Model.....	54
BAB 5 KESIMPULAN DAN SARAN	59
5.1 Kesimpulan	59
5.2 Saran.....	59
DAFTAR PUSTAKA	61
LAMPIRAN.....	64

DAFTAR TABEL

Tabel 2. 1 Daftar Penelitian Sebelumnya	10
Tabel 3. 1 <i>Confusion Matrix</i>	24
Tabel 4. 1 Hasil <i>Cleaning Data</i>	33
Tabel 4. 2 Hasil <i>Case Folding</i>	34
Tabel 4. 3 Hasil <i>Tokenizing</i>	35
Tabel 4. 4 Hasil <i>Normalization</i>	36
Tabel 4. 5 Hasil <i>Stopword Removal</i>	38
Tabel 4. 6 Hasil <i>Stemming</i>	39
Tabel 4. 7 Contoh Pelabelan Manual	41
Tabel 4. 8 Hasil Perhitungan TF-IDF	47
Tabel 4. 9 Hasil Rata-rata Akurasi Setiap Nilai k	52
Tabel 4. 10 Rangkuman Perhitungan <i>Confusion Matrix</i>	55
Tabel 4. 11 Hasil <i>Confusion Matrix</i>	57

DAFTAR GAMBAR

Gambar 3. 1 Alur Penelitian	19
Gambar 4. 1 Hasil Pengumpulan Data	29
Gambar 4. 2 Hasil Data <i>Review</i>	30
Gambar 4. 3 Hasil <i>Pre-processing</i>	40
Gambar 4. 4 Hasil <i>Labeling</i>	42
Gambar 4. 5 Hasil Visualisasi <i>Labeling</i>	43
Gambar 4. 6 Hasil Visualisasi <i>WordCloud</i>	44
Gambar 4. 7 Hasil Visualisasi <i>WordCloud</i>	45
Gambar 4. 8 Hasil Perhitungan TF-IDF	47
Gambar 4. 9 Hasil Visualisasi Kata Tertinggi	48
Gambar 4. 10 Hasil Visualisasi Rata-rata Akurasi Nilai k	53
Gambar 4. 11 Hasil Perhitungan <i>Confusion Matrix</i>	55
Gambar 4. 12 Hasil Visualisasi <i>Confusion Matrix</i>	56

DAFTAR LAMPIRAN

Lampiran 1 <i>Script</i> Pemrograman	64
Lampiran 2 Data <i>Cyberbullying</i>	67
Lampiran 3 Jadwal Penelitian	77
Lampiran 4 Lembar Bimbingan Dosen	78
Lampiran 5 Hasil Cek Plagiarisme	79

PERPUSTAKAAN
UNIVERSITAS JENDERAL ACHMAD YANI
YOGYAKARTA

DAFTAR SINGKATAN

AI	<i>Artificial Intelligence</i>
API	<i>Application Programming Interface</i>
BERT	<i>Bidirectional Encoder Representations from Transformers</i>
CSV	<i>Comma Separated Values</i>
FN	<i>False Negative</i>
FP	<i>False Positive</i>
KNN	<i>K-Nearest Neighbor</i>
LSTM	<i>Long Short-Term Memory</i>
ML	<i>Machine Learning</i>
NLP	<i>Natural Language Processing</i>
NLTK	<i>Natural Language Toolkit</i>
SVM	<i>Support Vector Machine</i>
TF-IDF	<i>Term Frequency - Inverse Document Frequency</i>
TN	<i>True Negative</i>
TP	<i>True Positive</i>