

TUGAS AKHIR

**ANALISIS SENTIMEN OPINI MASYARAKAT PADA
KOMENTAR YOUTUBE TERHADAP
PELAKSANAAN KTT G20 MENGGUNAKAN
METODE *SUPPORT VECTOR MACHINE* (SVM)**



TIARA DITA WARANGGANI

19103142

**PROGRAM STUDI S1 SISTEM INFORMASI
FAKULTAS INFORMATIKA
INSTITUT TEKNOLOGI TELKOM PURWOKERTO
2023**

TUGAS AKHIR

**ANALISIS SENTIMEN OPINI MASYARAKAT PADA
KOMENTAR YOUTUBE TERHADAP
PELAKSANAAN KTT G20 MENGGUNAKAN
METODE *SUPPORT VECTOR MACHINE* (SVM)**

**ANALYSIS OF PUBLIC OPINION SENTIMENT ON
YOUTUBE COMMENTS ON THE
IMPLEMENTATION OF THE G20 SUMMIT USING
THE SUPPORT VECTOR MACHINE (SVM) METHOD**

Disusun Sebagai Salah Satu Syarat Untuk Memperoleh Gelar Sarjana Komputer



TIARA DITA WARANGGANI

19103142

**PROGRAM STUDI S1 SISTEM INFORMASI
FAKULTAS INFORMATIKA
INSTITUT TEKNOLOGI TELKOM PURWOKERTO
2023**

LEMBAR PERSETUJUAN PEMBIMBING

LEMBAR PERSETUJUAN PEMBIMBING

ANALISIS SENTIMEN OPINI MASYARAKAT PADA KOMENTAR YOUTUBE TERHADAP PELAKSANAAN KTT G20 MENGGUNAKAN METODE *SUPPORT VECTOR MACHINE* (SVM)

ANALYSIS OF PUBLIC OPINION SENTIMENT ON YOUTUBE COMMENTS ON THE IMPLEMENTATION OF THE G20 SUMMIT USING THE SUPPORT VECTOR MACHINE (SVM) METHOD

Dipersiapkan dan Disusun Oleh

Tiara Dita Waranggani

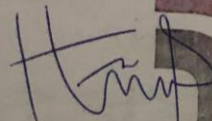
19103142

Fakultas Informatika

Institut Teknologi Telkom Purwokerto

Pada Tanggal: 07 Agustus 2023

Pembimbing Utama,



Khairun Nisa Meiah N., S.Pd., M.Kom

NIDN. 0618059401

LEMBAR PENGESAHAN

LEMBAR PENGESAHAN

**ANALISIS SENTIMEN OPINI MASYARAKAT PADA
KOMENTAR YOUTUBE TERHADAP PELAKSANAAN KTT
G20 MENGGUNAKAN METODE *SUPPORT VECTOR
MACHINE (SVM)***

***ANALYSIS OF PUBLIC OPINION SENTIMENT ON YOUTUBE
COMMENTS ON THE IMPLEMENTATION OF THE G20
SUMMIT USING THE SUPPORT VECTOR MACHINE (SVM)
METHOD***

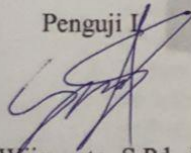
Disusun Oleh

Tiara Dita Waranggani 19103142

Telah Diujikan dan Dipertahankan dalam Sidang Ujian Tugas

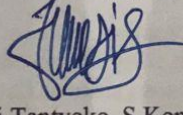
Akhir Pada 18 Oktober 2023

Penguji I,



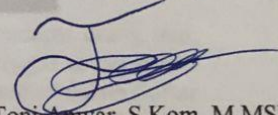
Sena Wijayanto, S.Pd., M.T.
NIDN.0613109201

Penguji II,



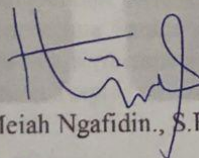
Henri Tantyoko, S.Kom., M.Kom.
NIDN.0627129601

Penguji III,



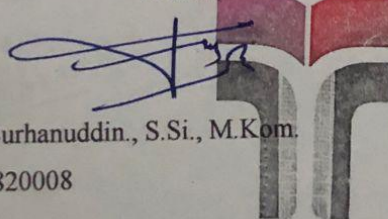
Tomi Anwar, S.Kom., M.MSI
NIDN.0613069102

Pembimbing Utama,



Khairun Nisa Meiah Ngafidin., S.Pd., M.Kom.
NIDN. 0618059401

Dekan,



Auliya Burhanuddin., S.Si., M.Kom.
NIK. 19820008

HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR

HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR

Yang bertandatangan di bawah ini,

Nama mahasiswa : Tiara Dita Waranggani

NIM : 19103142

Program Studi : Sistem Informasi

Menyatakan bahwa Tugas Akhir dengan judul berikut:

**ANALISIS SENTIMEN OPINI MASYARAKAT PADA KOMENTAR
YOUTUBE TERHADAP PELAKSANAAN KTT G20
MENGUNAKAN METODE *SUPPORT VECTOR MACHINE (SVM)***

Dosen Pembimbing Utama : Khoirun Nisa Meiah Ngafidin., S.Pd., M.Kom

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Institut Teknologi Telkom Purwokerto maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan, dan penelitian Saya Sendiri, tanpa bantuan pihak lain kecuali arahan dari Tim Dosen Pembimbing.
3. Dalam Karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggungjawab Saya, bukan tanggungjawab Institut Teknologi Telkom Purwokerto.
5. Pernyataan ini Saya buat dengan sesungguhnya, apabila dikemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka Saya bersedia menerima Sanksi Akademik dengan pencabutan gelar yang sudah diperoleh serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Purwokerto, 27 Agustus 2023,

Yang Menyatakan,



10000
METERAN
TEMPEL
24DB7AKX645511200

(Tiara Dita Waranggani)

KATA PENGANTAR

Segala puji dan syukur penulis panjatkan kehadirat Allah SWT atas limpahan rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan Tugas Akhir dan menyusun skripsi ini dengan lancar sebagai salah satu syarat untuk menyelesaikan Program Sarjana (S1) Jurusan Sistem Informasi Fakultas Informatika Institut Teknologi Telkom Purwokerto. Penulis menyadari bahwa skripsi ini tidak mungkin terselesaikan tanpa adanya dukungan, bantuan, bimbingan dan nasehat dari berbagai pihak selama penyusunan skripsi ini. Pada kesempatan ini penulis menyampaikan terima kasih setulus-tulusnya kepada:

1. Allah SWT yang telah memberikan ridho, nikmat, serta iman dalam penyusunan Tugas Akhir Skripsi.
2. Ibu Dr. Tenia Wahyuningrum, S.Kom., M.T. selaku Rektor Institut Teknologi Telkom Purwokerto.
3. Bapak Auliya Burhanuddin, S.Si., M.Kom. selaku Dekan Fakultas Informatika Intitut Teknologi Telkom Purwokerto.
4. Ibu Dwi Mustika Kusumawardani, S.Kom., M.Kom. selaku Ketua Program Studi Sistem Informasi, sekaligus dosen wali yang telah membantu penulis dalam mengikuti dan menyelesaikan studi di Fakultas Informatika, Institut Teknologi Telkom Purwokerto.
5. Ibu Khairun Nisa Meiah Ngafidin, S.Pd., M.Kom. selaku Pembimbing Utama yang sudah sabar dan turut memberikan solusi selama proses penyelesaian Tugas Akhir.
6. Bapak/Ibu dosen Fakultas Informatika yang telah memberikan pengetahuan selama perkuliahan.
7. Kedua orang tua penulis, yang selalu memberikan kasih sayang, doa, nasehat, serta atas kesabarannya yang luar biasa dalam setiap langkah hidup penulis, yang merupakan anugerah terbesar dalam hidup, Penulis berharap dapat menjadi anak yang dapat dibanggakan.
8. Seluruh teman-teman, sahabat, saudara-saudaridan semua pihak yang tidak dapat disebutkan satu per satu yang telah memberikan motivasi sehingga dapat terselesaikannya Tugas Akhir ini.

Dalam penulisan skripsi ini masih banyak kekurangan dan kesalahan, karena itu segala kritik dan saran yang membangun akan menyempurnakan penulisan skripsi ini serta bermanfaat bagi penulis dan para pembaca.

Purwokerto, 27 Oktober 2023

A handwritten signature in black ink, appearing to read 'Tiara Dita Waranggani', with a large, sweeping flourish above the name.

(Tiara Dita Waranggani)

DAFTAR ISI

LEMBAR PERSETUJUAN PEMBIMBING	iii
LEMBAR PENGESAHAN	iv
HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR	v
KATA PENGANTAR	vi
DAFTAR ISI.....	viii
DAFTAR TABEL.....	xi
DAFTAR GAMBAR	xii
DAFTAR SINGKATAN	xiii
DAFTAR LAMPIRAN.....	xiv
ABSTRAK.....	xv
ABSTRACT.....	xvi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Pertanyaan Penelitian	5
1.4 Batasan Masalah.....	5
1.5 Tujuan Penelitian.....	5
1.6 Manfaat Penelitian.....	6

BAB II TINJAUAN PUSTAKA DAN LANDASAN TEORI.....	7
2.1 Tinjauan Pustaka	7
2.2 Dasar Teori	20
2.2.1 Data Mining.....	20
2.2.2 API	20
2.2.3 Analisis Sentimen.....	20
2.2.4 Text Preprocessing	20
2.2.5 TF-IDF	21
2.2.6 Sastrawi	22
2.2.7 SentiWordNet.....	22
2.2.8 Support Vector Machine (SVM).....	23
2.2.9 Confusion Matrix	25
BAB III METODOLOGI PENELITIAN	29
3.1 Subjek dan Objek Penelitian	29
3.2 Alat dan Bahan Penelitian	29
3.3 Diagram Alir Penelitian	30
3.3.1 Identifikasi Masalah	31
3.3.2 Studi Literatur	31
3.3.3 Perumusan Masalah dan Tujuan Penelitian	31
3.3.4 Pengumpulan dan Analisis Data.....	32

3.3.5	Pengolahan Data.....	33
3.3.6	Analisa Hasil Algoritma Support Vector Machine	34
3.3.7	Evaluasi Hasil Analisis Sentimen.....	37
BAB IV HASIL DAN PEMBAHASAN		38
4.1	Pengumpulan dan Analisis Data.....	38
4.2	Pengolahan Data.....	42
4.2.1	Data Cleansing	44
4.2.2	Case Folding.....	46
4.2.3	Normalization.....	47
4.2.4	Tokenization.....	48
4.2.5	Stopword Removal.....	49
4.2.6	Stemming	50
4.2.7	TF-IDF	51
4.3	Analisis dengan SVM.....	51
4.4	Evaluasi Hasil analisis Sentimen.....	54
BAB V KESIMPULAN DAN SARAN.....		63
5.1	Kesimpulan.....	63
5.2	Saran.....	63
DAFTAR PUSTAKA		64
LAMPIRAN.....		70

DAFTAR TABEL

Tabel 2.1 Kajian Penelitian Terdahulu	10
Tabel 2.2 <i>Confusion matrix 2x2</i>	26
Tabel 3.1 Contoh Perhitungan SVM.....	35
Tabel 4.1 Tabel Data teratas yang didapatkan	43
Tabel 4.2 Contoh Data sebelum dilakukan <i>Cleansing</i>	44
Tabel 4.3 Hasil Data cleansing	45
Tabel 4.4 Hasil <i>Casefolding</i>	47
Tabel 4.5 Hasil Normalisasi.....	48
Tabel 4. 6 Hasil Tokenisasi.....	48
Tabel 4. 7 Hasil <i>stopword</i>	50
Tabel 4. 8 Hasil <i>Stemmer</i>	51
Tabel 4. 9 Pembobotan kata.....	51
Tabel 4. 10 Parameter Terbaik SVM	52
Tabel 4. 11 Perbandingan Akurasi prediksi 80:20.....	53
Tabel 4. 12 Perbandingan Akurasi prediksi 70:30.....	53
Tabel 4. 13 Perbandingan Akurasi prediksi 75:25	54
Tabel 4. 14 Evaluasi Kernel Linear prediksi 80:20.....	55
Tabel 4. 15 Evaluasi Kernel Linear skenario 70:30.....	56
Tabel 4. 16 Evaluasi Kernel Linear skenario 75:25.....	56
Tabel 4. 17 Evaluasi Kernel Sigmoid skenario 80:20.....	57
Tabel 4. 18 Evaluasi Kernel Sigmoid skenario 70:30.....	57
Tabel 4. 19 Evaluasi Kernel Sigmoid skenario 75:25.....	58
Tabel 4. 20 Evaluasi Kernel Polynomial skenario 80:20	58
Tabel 4. 21 Evaluasi Kernel Polynomial skenario 70:30	59
Tabel 4. 22 Evaluasi Kernel Polynomial skenario 75:25	59
Tabel 4. 23 Evaluasi Kernel RBF skenario 80:20.....	60
Tabel 4. 24 Evaluasi Kernel RBF skenario 70:30.....	60
Tabel 4. 25 Evaluasi Kernel RBF skenario 75:25.....	61
Tabel 4. 26 Evaluasi Model <i>Confussion matrix</i>	61

DAFTAR GAMBAR

Gambar 2. 1 Visualisasi klasifikasi SVM	24
Gambar 2. 2 Hyperplane yang memisahkan dua kelas positif(+1) dan negatif(-1)	24
Gambar 3. 1 Proses Penelitian	30
Gambar 3. 2 Diagram alir pengumpulan data	32
Gambar 3. 3 Analisis Data	33
Gambar 3. 4 Tahap <i>pre-processing</i>	33
Gambar 4. 1 Youtube Data API v3	38
Gambar 4.2 Potongan Cell Fungsi Crawling	39
Gambar 4.3 Proses <i>Crawling</i>	40
Gambar 4.4 Mengubah hasil Crawling ke Dataframe	40
Gambar 4.5 Menyimpan hasil dataset kedalam format csv	41
Gambar 4. 6 Analisis Data	41
Gambar 4. 7 Worldcloud Analysis	42
Gambar 4.8 Proses menghapus data kosong	45
Gambar 4.9 Proses membersihkan teks	45
Gambar 4.10 Total data setelah <i>data cleansing</i>	46
Gambar 4. 11 Proses <i>casefolding</i>	46
Gambar 4. 12 Fungsi Normalisasi	47
Gambar 4. 13 Proses Tokenisasi	48
Gambar 4. 14 Modul stopwords dan sastrawi	49
Gambar 4. 15 Tahap Stopword	49
Gambar 4. 16 Module <i>Stemmer</i>	50
Gambar 4. 17 Persentase dua kelas sentimen	54

DAFTAR SINGKATAN

SVM	= Support Vector Machine
NB	= Naïve Bayes
DF	= Decision Tree
RF	= Random Forest
TF-IDF	= Term Frequency-Inverse Document Frequency
K-NN	= K-Nearest Neighbor
KTT	= Konferensi Tingkat Tinggi
API	= Application Programming Interface
RBF	= Radial Basis Function

DAFTAR LAMPIRAN

Lampiran 1 Source Code Pre-processing.....	70
Lampiran 2 Source Code SVM.....	75
Lampiran 3 Cek Plagiasi Turnitin.....	80