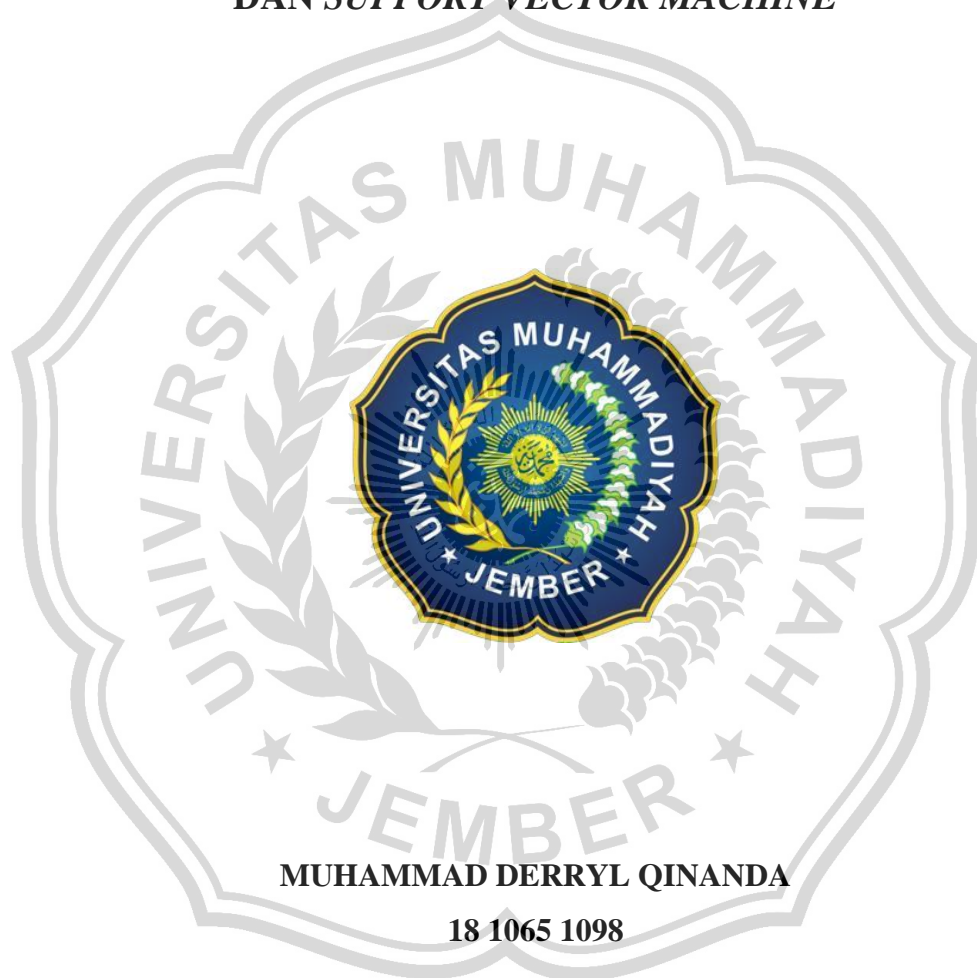


TUGAS AKHIR

**AKURASI SENTIMEN KOMENTAR YOUTUBE TENTANG
PENCEGAHAN DAN PENANGANAN KEKERASAN
SEKSUAL PADA PERMENDIKBUD BERBASIS *NAÏVE BAYES*
DAN *SUPPORT VECTOR MACHINE***



MUHAMMAD DERRYL QINANDA

18 1065 1098

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

2022

TUGAS AKHIR

**AKURASI SENTIMEN KOMENTAR YOUTUBE TENTANG
PENCEGAHAN DAN PENANGANAN KEKERASAN
SEKSUAL PADA PERMENDIKBUD BERBASIS *NAÏVE BAYES*
DAN *SUPPORT VECTOR MACHINE***

Disusun Untuk Melengkapi dan Memenuhi Syarat Kelulusan
Guna Meraih Sarjana Komputer
Teknik Informatika Universitas Muhammadiyah Jember



MUHAMMAD DERRYL QINANDA

1810651098

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

2022

HALAMAN PERSETUJUAN SIDANG TUGAS AKHIR

**AKURASI SENTIMEN KOMENTAR YOUTUBE TENTANG
PENCEGAHAN DAN PENANGANAN KEKERASAN
SEKSUAL PADA PERMENDIKBUD BERBASIS *NAÏVE BAYES*
DAN *SUPPORT VECTOR MACHINE***

Oleh:

Muhammad Derryl Qinanda

1810651098

Telah disetujui bahwa 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



Agung Nilogiri, S.T., M.Kom.

NIDN. 0030037701

Pembimbing II



Taufiq Timur W., S.Kom., M.Kom.

NIDN. 0705078006

HALAMAN PENGESAHAN
AKURASI SENTIMEN PADA KOMENTAR YOUTUBE
UNTUK PERMENDIKBUD TENTANG PENCEGAHAN DAN
PENANGANAN KEKERASAN SEKSUAL BERBASIS NAÏVE
BAYES DAN SUPPORT VECTOR MACHINE

Oleh:

Muhammad Derryl Qinanda
1810651098

Telah mempertanggung jawabkan Laporan Tugas Akhirnya pada sidang Tugas Akhir tanggal 29 Juni 2022 sebagai salah satu syarat kelulusan dan mendapatkan gelar Sarjana Komputer (S.Kom)

di
Universitas Muhammadiyah Jember

Disetujui oleh,

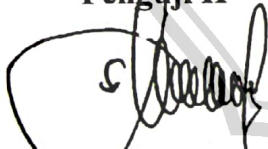
Penguji I



Wiwik Suharso, S.Kom, M.Kom

NIDN. 0006097601

Penguji II



Ilham Saifudin, S.Pd., M.Si

NIDN. 0731108903

Mengesahkan,

Dekan Fakultas Teknik



Drs. Nanang Saiful Rizal, ST., MT., IPM

NIDN. 0705047806

Pembimbing I



Agung Nilogiri, S.T., M.Kom.

NIDN. 0030037701

Pembimbing II



Taufiq Timur W., S.Kom., M.Kom.

NIDN. 0705078006

Mengetahui,

Ketua Program Studi Teknik Informatika



Ari Eko Wardono, S.Kom, M.Kom

NIDN. 0014027501

PERNYATAAN

Yang bertanda tangan di bawah ini:

Nama : Muhammad Derryl Qinanda

NIM : 1810651098

Program Studi : Teknik Informatika

Perguruan Tinggi : Universitas Muhammadiyah Jember

Menyatakan bahwa Tugas Akhir yang berjudul **“Akurasi Sentimen Komentar Youtube Tentang Pencegahan Dan Penanganan Kekerasan Seksual Pada Permendikbud Berbasis Naïve Bayes Dan Support Vector Machine”** adalah ASLI dan BELUM PERNAH dibuat orang lain, kecuali yang diacu dalam Daftar Pustaka pada Tugas Akhir ini.

Demikian Surat Pernyataan ini dibuat dengan sebenar-benarnya dan apabila pernyataan ini tidak sesuai, penulis bersedia mendapat sanksi dari akademik.

Jember, 29 Juni 2022



Muhammad Derryl Qinanda

KATA PENGANTAR

Alhamdulillah puji Allah atas segala yang telah dilimpahkan serta karunianya yang sangat luar biasa dalam hidup ini kepada penulis berupa nikmat untuk selalu bersujud dan meminta ampunan kepada-Nya. Serta berkat ramat-Nya penulis ini diberikan kemudahan dalam menyelesaikan studi di kampus Universitas Muhammadiyah Jember.

Atas segala upaya, bimbingan, dan arahan dari semua pihak, penulis mengucapkan banyak terimakasih. Dalam kesempatan ini penulis mengucapkan terima kasih sebanyak-banyaknya kepada:

1. Allah SWT atas segala rahmat, nikmat, dan hidayah-Nya, terima kasih atas kehendak-Mu telah memberikan gelar Sarjana Komputer.
2. Bapak Ari Eko Wardoyo, M. Kom. selaku Ketua Program Studi Teknik Informatika Universitas Muhammadiyah Jember.
3. Bapak Agung Nilogiri, S.T., M.Kom. selaku dosen pembimbing I dan Bapak Taufiq Timur W., S.Kom., M.Kom. selaku dosen pembimbing II yang telah meluangkan waktu untuk memberikan arahan dan membimbing penulis dalam menyelesaikan tugas akhir ini.
4. Bapak Wiwik Suharso, S.Kom, M.Kom. selaku dosen penguji I dan Bapak Ilham Saifudin, S.Pd., M.Si. selaku dosen penguji II yang telah memberikan saran dan masukan dalam penyelesaian tugas akhir ini.
5. Kepada Bapak dan Ibu Dosen Program Studi Teknik Informatika Universitas Muhammadiyah Jember yang telah memberikan ilmunya kepada penulis.
6. Kepada kedua orang tua dan adik yang selalu mendoakan, memberi nasihat dan dukungan sehingga penulis dapat menyelesaikan tugas akhir ini.
7. Kepada Salsabila Mazya Permataning Tyas. S.Kom. yang selalu senantiasa ada, memberikan semangat, dukungan dan selalu sangat sabar selama menemani penulis.

8. Kepada Genta Sufianto teman seperjuangan yang selalu mendukung satu sama lain dalam mencapai target kuliah.
9. Kepada Dimas, Rusli, Nanda, Arifin, Firman dan Alif teman seperjuangan belajar bersama mulai dari awal kuliah sampai akhir. Serta untuk teman-teman mahasiswa Teknik Informatika Angkatan 2018 yang sudah kebersamai dalam masa studi di Teknik Informatika Universitas Muhammadiyah Jember.
10. Kepada teman-teman UPT-Pusat Data dan Informasi yang telah memberikan dukungan untuk menyelesaikan Tugas Akhir ini.

Penulis menyadari bahwa dalam penyusunan laporan ini masih banyak kekurangan, oleh karena itu kritik dan saran yang bersifat membangun akan sangat berguna kearah kesempurnaan dan semoga tugas akhir ini dapat berguna bagi rekan-rekan mahasiswa fakultas Teknik pada khususnya Teknik informatika dan pembaca pada umumnya.

Dan untuk semuanya yang tersebut diatas, semoga Allah SWT menunjukkan kita pada jalan yang benar, menghimpunkan kita dengan orang-orang yang beriman dan berilmu, diampuni dosa-dosa kita dan senantiasa tercurahkan rahmat, taufiq dan hidayah-Nya kepada kita dan keluarga kita sekalian. Aamiin.

Motto

“Pengetahuan tidak hanya didasarkan pada kebenaran saja
Tetapi juga Kesalahan”
(Carl Gustav Jung)

Jangan pernah berhenti berharap, karena Allah lebih tau saat yang tepat
mengabulkan permintaan.



DAFTAR ISI

HALAMAN PERSETUJUAN.....	i
HALAMAN PENGESAHAN.....	ii
HALAMAN PERNYATAAN.....	iii
ABSTRAK	iv
ABSTRACT	v
KATA PENGANTAR.....	vi
MOTTO	viii
DAFTAR ISI.....	ix
DAFTAR GAMBAR.....	xi
DAFTAR TABEL.....	xii
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 Masalah.....	4
BAB II TINJAUAN PUSTAKA.....	5
2.1 Text Mining.....	5
2.2 Analisis Sentimen.....	6
2.3 Youtube.....	6
2.4 TF - IDF.....	7
2.5 Metode Klasifikasi.....	8
2.5.1 Support Vector Machine.....	8
2.5.2 Gaussian Naïve Bayes.....	10
2.6 K – Fold Cross Validation.....	11
2.7 Python.....	12
2.8 Penelitian Terdahulu.....	13
BAB III METODOLOGI PENELITIAN.....	14
3.1 Flowchart Tahapan Penelitian.....	14

3.2	Pengambilan Data	14
3.3	Labelling Data	15
3.4	Pembagian Data	16
3.5	Preprocessing	16
3.6	TF – IDF	21
3.7	K – Fold Validation	24
3.8	Support Vector Machine	27
3.9	Gaussian Naïve Bayes	32
BAB IV	HASIL DAN PEMBAHASAN	36
4.1	Implementasi Analisis Sentimen	36
4.1.1	<i>Scrapping Data Komentar Youtube</i>	36
4.1.2	<i>Labelling</i>	37
4.1.3	<i>Pembagian Data</i>	38
4.1.4	<i>Preprocessing Text</i>	39
4.1.5	Pembobotan Kata dengan TF-IDF.....	43
4.1.6	<i>Support Vector Machine</i>	47
4.1.7	<i>Gaussian Naïve Bayes</i>	52
4.1.8	<i>K – Fold Cross Validation</i>	56
4.1.9	Pengujian Data dengan <i>Unseen Data Test</i>	62
4.2	Analisis Hasil	67
BAB V	KESIMPULAN DAN SARAN	70
5.1	Kesimpulan	70
5.2	Saran	70
	DAFTAR PUSTAKA	71
	LAMPIRAN-LAMPIRAN	73

DAFTAR GAMBAR

Gambar 2.1	Contoh Model 3 – <i>Fold Cross Validation</i>	12
Gambar 3.1	Flowchart Penelitian	14
Gambar 4.1	Proses scrapping data pada komentar youtube	34
Gambar 4.2	Hasil scrapping data dengan format csv	34
Gambar 4.3	Data dengan label negatif	35
Gambar 4.4	Data dengan label positif	35
Gambar 4.5	Grafik Tingkat Akurasi <i>Naïve Bayes</i> K=2.....	57
Gambar 4.6	Grafik Tingkat Akurasi <i>Naïve Bayes</i> K=5.....	57
Gambar 4.7	Grafik Tingkat Akurasi <i>Naïve Bayes</i> K=10.....	58
Gambar 4.8	Grafik Tingkat Akurasi <i>Support Vector Machine</i> K=2	60
Gambar 4.9	Grafik Tingkat Akurasi <i>Support Vector Machine</i> K=5	60
Gambar 4.10	Grafik Tingkat Akurasi <i>Support Vector Machine</i> K=10	61
Gambar 4.11	Grafik Perbandingan Tingkat Akurasi <i>Naïve Bayes</i> K=2.....	64
Gambar 4.12	Grafik Perbandingan Tingkat Akurasi <i>Naïve Bayes</i> K=5.....	64
Gambar 4.13	Grafik Perbandingan Tingkat Akurasi <i>Naïve Bayes</i> K=10.....	64
Gambar 4.14	Grafik Perbandingan Tingkat Akurasi <i>Support Vector Machine</i> K=2	66
Gambar 4.15	Grafik Perbandingan Tingkat Akurasi <i>Support Vector Machine</i> K=5	66
Gambar 4.16	Grafik Perbandingan Tingkat Akurasi <i>Support Vector Machine</i> K=10	66

DAFTAR TABEL

Tabel 2.1	Penelitian Terdahulu	13
Tabel 3.1	Contoh Data Komentar Hasil <i>Scrapping</i>	15
Tabel 3.2	Contoh Pelabelan Data Komentar	15
Tabel 3.3	Tahap <i>Cleansing</i>	17
Tabel 3.4	Tahap <i>Case Folding</i>	18
Tabel 3.5	Tahap <i>Tokenizing</i>	19
Tabel 3.6	Tahap <i>Stopword Removal</i>	20
Tabel 3.7	Tahap <i>Stemming</i>	21
Tabel 3.8	Contoh dokumen yang akan dihitung nilai TF-IDF	21
Tabel 3.9	Perhitungan TF-IDF	22
Tabel 3.10	Nilai TF.IDF pada dokumen	23
Tabel 3.11	Sampel Data yang di Vektorisasi	28
Tabel 3.12	Tabel Fungsi Kernel.....	28
Tabel 3.13	Hasil Fungsi Kernel.....	29
Tabel 3.14	Tabel Matriks	29
Tabel 3.15	Hasil Perhitungan Nilai Error.....	30
Tabel 3.16	Hasil Perhitungan Delta Alpha.....	30
Tabel 3.17	Nilai Alpha Baru	30
Tabel 3.18	Hasil Nilai Perhitungan Sentimen.....	31
Tabel 3.19	Contoh Data Bobot TF-IDF	32
Tabel 3.20	Jumlah Kelas	32
Tabel 3.21	Nilai Rata-Rata (<i>Mean</i>)	33
Tabel 3.22	Standar Deviasi	33
Tabel 3.23	<i>Gaussian</i> di Setiap Dokumen.....	34
Tabel 4.1	Hasil Labelling	38
Tabel 4.2	Hasil pembagian data Training dan Unseen.....	39
Tabel 4.3	Tahap Preprocessing Text	39
Tabel 4.4	Perhitungan TF-IDF	43
Tabel 4.5	Nilai TF-IDF pada dokumen.....	45

Tabel 4.6	Tabel TF-IDF	46
Tabel 4.7	Sampel Data yang di Vektorisasi	48
Tabel 4.8	Tabel Fungsi Kernel.....	48
Tabel 4.9	Hasil Fungsi Kernel.....	49
Tabel 4.10	Tabel Matriks	49
Tabel 4.11	Hasil Perhitungan Nilai Error.....	50
Tabel 4.12	Hasil Perhitungan Delta Alpha.....	50
Tabel 4.13	Nilai Alpha Baru	51
Tabel 4.14	Hasil Nilai Perhitungan Sentimen.....	52
Tabel 4.15	Contoh data bobot TF-IDf.....	52
Tabel 4.16	Jumlah Kelas	53
Tabel 4.17	Nilai Rata-Rata (<i>Mean</i>)	53
Tabel 4.18	Standar deviasi	54
Tabel 4.19	<i>Gaussian</i> di Setiap Dokumen.....	54
Tabel 4.20	Hasil Prediksi	55
Tabel 4.21	Nilai Akurasi Metode <i>Naïve Bayes</i>	59
Tabel 4.22	Nilai Akurasi Metode <i>Support Vector Machine</i>	62
Tabel 4.23	Pengujian <i>Unseen Data Test Naïve Bayes</i>	63
Tabel 4.24	Pengujian <i>Unseen Data Test Support Vector Machine</i>	65
Tabel 4.25	Hasil Perhitungan pada metode <i>Naïve Bayes</i> dan <i>Support Vector Machine</i>	67
Tabel 4.26	Hasil Perhitungan pada metode <i>Naïve Bayes</i> dan <i>Support Vector Machine - 2</i>	68