

TUGAS AKHIR

ANALISIS SENTIMEN KOMENTAR *YOUTUBE* TERKAIT KENAIKAN PPN 12% DENGAN *MULTINOMIAL NAÏVE* *BAYES*



TRIA AGUSTINA

2110651005

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS TEKNIK

UNIVERSITAS MUHAMMADIYAH JEMBER

2025

TUGAS AKHIR

ANALISIS SENTIMEN KOMETAR *YOUTUBE* TERKAIT KENAIKAN PPN 12% DENGAN *MULTINOMIAL NAÏVE* *BAYES*

Disusun untuk Memenuhi Persyaratan Guna Meraih Gelar Sarjana
Komputer Teknik Informatika Universitas Muhammadiyah
Jember



TRIA AGUSTINA
2110651005

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH JEMBER

2025

HALAMAN PERSETUJUAN SIDANG TUGAS AKHIR
ANALISIS SENTIMEN KOMENTAR *YOUTUBE* TERKAIT KENAIKAN
PPN 12% DENGAN *MULTINOMIAL NAÏVE BAYES*

Oleh :

Tria Agustina

2110651005

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

Pembimbing II


Hardian Oktavianto, S.Si., M. Kom
NIDN.0722108105


Wiwik Suharso, M. Kom
NIDN.0006097601

HALAMAN PENGESAHAN

ANALISIS SENTIMEN KOMENTAR *YOUTUBE* TERKAIT KENAIKAN PPN 12% DENGAN *MULTINOMIAL NAÏVE BAYES*

Oleh :

Tria Agustina

2110651005

Telah mempertanggungjawabkan Laporan Tugas Akhir pada Sidang Tugas Akhir
9 Agustus 2025 sebagai salah satu syarat kelulusan dan
mendapatkan gelar sarjana Komputer (S.Kom)

di

Universitas Muhammadiyah Jember

Disetujui Oleh,

Penguji I



Dudi Irawan, S. T., M. Kom
NIDN.0730037703

Pembimbing I



Hardian Oktavianto, S.Si., M. Kom
NIDN.0722108105

Penguji II



Dr. Dewi Lusiana Pater, MT
NIDN.0712086702

Pembimbing II



Wiwik Suharso, M. Kom
NIDN.0006097601

Dekan Fakultas Teknik



Dr. Ir. Muhtar, S.T., M.T., IPM
NIDN.0010067301

Ketua Program Studi Teknik
Informatika



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

HALAMAN PERNYATAAN

Yang bertanda tangan di bawah ini :

Nama : Tria Agustina
NIM : 2110651005
Program Studi : Teknik Informatika
Perguruan Tinggi : Universitas Muhammadiyah Jember

Menyatakan bahwa Tugas Akhir yang berjudul "**ANALISIS SENTIMEN KOMENTAR *YOUTUBE* TERKAIT KENAIKAN PPN 12% DENGAN *MULTINOMIAL NAÏVE BAYES***" bukan merupakan Tugas Akhir orang lain baik sebagian maupun keseluruhan kecuali dalam bentuk kutipan yang telah disebutkan sumbernya.

Demikian Surat pernyataan ini dibuat dengan sebenar-benarnya dan apabila pernyataan ini tidak benar penulis bersedia mendapatkan sanksi akademik.

Jember, 9 Agustus 2025

Menyatakan



Tria Agustina
2110651005

KATA PENGANTAR

Puji dan syukur penulis panjatkan ke hadirat Allah SWT atas segala rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan dapat menyelesaikan penelitian dan penyusunan laporan yang berjudul "*Analisis Sentimen Komentar YouTube terhadap Kenaikan PPN 12% Dengan Multinomial Naïve Bayes*" dengan baik dan tepat waktu.

Penulis sangat berterima kasih kepada Bapak Hardian Oktavianto, S.Kom., M.Kom. selaku pembimbing pertama dan Bapak Wiwik Suharso, M. Kom. selaku pembimbing kedua, atas segala perhatian dan bimbingannya serta arahan-arahan yang diberikan kepada penulis dalam upaya menyelesaikan tugas akhir ini.

Ucapan terimakasih juga penulis sampaikan kepada penguji 1 Bapak Dudi Irawan, S.T., M.Kom dan penguji 2 Ibu Ir.Dewi Lusiana Pater, MT atas arahan dan masukannya sampai sidang tugas akhir ini.

Ucapan terima kasih juga disampaikan penulis kepada Dr. Ir. Muhtar, S.T., M.T., IPM. selaku Dekan Fakultas Teknik, atas kesediaannya penulis belajar di Fakultas Teknik Prodi Teknik Informatika Universitas Muhammadiyah Jember.

Penulis menyadari bahwa laporan ini masih jauh dari kata sempurna. Oleh karena itu, saran dan kritik yang membangun sangat diharapkan guna perbaikan di masa mendatang. Demikian kata pengantar dari penulis, Semoga dapat memberikan manfaat bagi pembaca.

DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN PERSETUJUAN SIDANG TUGAS AKHIR.....	ii
HALAMAN PENGESAHAN.....	iii
HALAMAN PERNYATAAN	iv
ABSTRAK.....	v
ABSTRACT.....	vi
PERSEMBAHAN	vii
MOTTO	ix
KATA PENGANTAR.....	x
DAFTAR ISI.....	xi
DAFTAR GAMBAR	xiii
DAFTAR TABEL.....	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah	4
1.3 Tujuan Penelitian	4
1.4 Manfaat Penelitian.....	4
1.5 Batasan Penelitian	4
BAB II TINJAUAN PUSTAKA.....	5
2.1 Penelitian terdahulu	5
2.2 Analisis Sentimen	7
2.3 <i>Text Mining</i>	7
2.4 <i>Labelling</i>	8
2.5 <i>Preprocessing Text</i>	8
2.6 <i>TF-IDF</i>	10
2.7 <i>Multinomial Naïve Bayes</i>	10
2.8 <i>Oversampling</i>	12
2.9 <i>K-Fold Cross Validation</i>	12
2.10 <i>Confusion Matrix</i>	13
2.11 <i>Python</i>	14
2.11.1 <i>Natural Language Toolkit (NLTK)</i>	15
2.11.2 <i>Sastrawi</i>	15
2.11.3 <i>Scikit-learn</i>	15
2.12 <i>Youtube</i>	15
2.13 <i>Google Colab</i>	16
BAB III METODOLOGI PENELITIAN.....	17
3.1 Tahapan penelitian.....	17
3.2 Pengumpulan Data.....	17
3.3 Pelabelan Data	18
3.4 <i>Preprocessing text</i>	19
3.5 Pembobotan <i>TF-IDF</i>	23
3.6 <i>Random Oversampling</i>	28
3.7 <i>K-Fold Cross Validation</i>	29
3.8 <i>Multinomial Naïve Bayes</i>	31
3.9 <i>Confusion Matrix</i>	33
BAB IV HASIL DAN PEMBAHASAN	35

4.1 Pengumpulan Data.....	35
4.2 <i>Labelling Data</i>	35
4.3 <i>Preprocessing Text</i>	36
4.4 Pembobotan <i>TF-IDF</i>	39
4.5 <i>Random Oversampling</i>	41
4.6 Pembagian dan Pengujian.....	41
4.7 Hasil Evaluasi	54
4.7.1 Tanpa <i>Oversampling</i>	54
4.7.2 <i>Oversampling</i>	56
4.8 Hasil Analisis.....	57
BAB V PENUTUP.....	60
5.1 Kesimpulan	60
5.2 Saran	60
DAFTAR PUSTAKA	61
LAMPIRAN.....	64



DAFTAR GAMBAR

Gambar 1. 1 Platform Sosial di Dunia (Sumber : simon kemp, 2024)	1
Gambar 2. 1 Ilustrasi Cross Validation (Sumber : https://scikit-learn.org/stable/modules/cross_validation.html#grid-search)	13
Gambar 4. 1 Hasil Pengumpulan Data	35
Gambar 4. 2 Hasil Pelabelan Data	36
Gambar 4. 3 Hasil Cleaning	36
Gambar 4. 4 Hasil Case Folding	37
Gambar 4. 5 Hasil Tokenisasi	37
Gambar 4. 6 Hasil Normalisasi	38
Gambar 4. 7 Hasil Stopword Removal.....	38
Gambar 4. 8 Hasil Stemming	39
Gambar 4. 9 Distribusi Data Sebelum dan Setelah Oversampling	41
Gambar 4. 10 Grafik Akurasi, Presisi, Recall 2-fold	42
Gambar 4. 11 Grafik Akurasi, Presisi, Recall 3-fold	43
Gambar 4. 12 Grafik Akurasi, Presisi, Recall 4-fold	44
Gambar 4. 13 Grafik Akurasi, Presisi, Recall 5-fold	46
Gambar 4. 14 Grafik Akurasi, Presisi, Recall 6-fold	47
Gambar 4. 15 Grafik Akurasi, Presisi, Recall 7-fold	49
Gambar 4. 16 Grafik Akurasi, Presisi, Recall 8-fold	50
Gambar 4. 17 Grafik Akurasi, Presisi, Recall 9-fold	51
Gambar 4. 18 Grafik Akurasi, Presisi, Recall 10-fold	53
Gambar 4. 19 Grafik Hasil Akurasi	58
Gambar 4. 20 Grafik Hasil Presisi	58
Gambar 4. 21 Grafik Hasil Recall	58
Gambar 4. 22 Rata-rata Akurasi per K.....	59
Gambar 4. 23 Rata-rata Presisi per K.....	59
Gambar 4. 24 Rata-rata Recall per K.....	59

DAFTAR TABEL

Tabel 2. 1 Penelitian Terdahulu	5
Tabel 2. 3 <i>Confusion Matrix</i>	14
Tabel 2. 4 Rumus Akurasi, Presisi, Recall	14
Tabel 3. 1 Contoh Hasil Crawling Komentar Youtube	18
Tabel 3. 2 Contoh Pelabelan	18
Tabel 3. 3 Contoh Tahap Cleansing	19
Tabel 3. 4 Contoh Tahap Case Folding	20
Tabel 3. 5 Contoh Tahap Tokenisasi	20
Tabel 3. 6 Contoh Tahap Normalisasi	21
Tabel 3. 7 Contoh Tahap Stopword Removal	22
Tabel 3. 8 Contoh Tahap Stemming	23
Tabel 3. 9 Contoh dokumen yang akan dihitung TF-IDF	24
Tabel 3. 10 Contoh Perhitungan TF	25
Tabel 3. 11 Contoh Perhitungan IDF	26
Tabel 3. 12 Contoh Perhitungan TF-IDF	27
Tabel 3. 13 Ilustrasi Random Oversampling	28
Tabel 3. 14 Contoh Nilai Probabilitas D3	32
Tabel 3. 15 Hasil Klasifikasi Sentimen	33
Tabel 3. 16 <i>Confusion Matrix</i>	34
Tabel 3. 17 Hasil Kriteria	34
Tabel 4. 1 Hasil TF dan IDF	39
Tabel 4. 2 Hasil TF-IDF	40
Tabel 4. 3 Hasil Pengujian 2-Fold	42
Tabel 4. 4 <i>Confusion Matrix</i> K-2 Langkah Uji 2	42
Tabel 4. 5 Hasil Pengujian 3-Fold	43
Tabel 4. 6 <i>Confusion Matrix</i> K-3 Langkah Uji 1	43
Tabel 4. 7 Hasil Pengujian 4-Fold	44
Tabel 4. 8 <i>Confusion Matrix</i> K-4 Langkah Uji 1	45
Tabel 4. 9 Hasil Pengujian 5-Fold	45
Tabel 4. 10 <i>Confusion Matrix</i> K-5 Langkah Uji 1	46
Tabel 4. 11 Hasil Pengujian 6-Fold	47
Tabel 4. 12 <i>Confusion Matrix</i> K-6 Langkah Uji 2	47
Tabel 4. 13 Hasil Pengujian 7-Fold	48
Tabel 4. 14 <i>Confusion Matrix</i> K-7 Langkah Uji 1	49
Tabel 4. 15 Hasil Pengujian 8-Fold	50
Tabel 4. 16 <i>Confusion Matrix</i> K-8 Langkah Uji 2	50
Tabel 4. 17 Hasil Pengujian 9-Fold	51
Tabel 4. 18 <i>Confusion Matrix</i> K-9 Langkah Uji 2	52
Tabel 4. 19 Hasil Pengujian 10-Fold	53
Tabel 4. 20 <i>Confusion Matrix</i> K-10 Langkah Uji 2	53
Tabel 4. 21 Rekapitulasi Akurasi, Presisi, Recall tanpa Oversampling	54
Tabel 4. 22 Rekapitulasi Akurasi, Presisi, Recall dengan Oversampling	56
Tabel 4. 23 Hasil Analisis	57