

TUGAS AKHIR

**SENTIMEN ANALISIS ULASAN DUOLINGO DENGAN
METODE ALGORITMA *MULTINOMIAL NAÏVE BAYES***



Robbi Zulcharnain

2010651076

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS TEKNIK

UNIVERSITAS MUHAMMADIYAH JEMBER

2025

TUGAS AKHIR

SENTIMEN ANALISIS ULASAN DUOLINGO DENGAN METODE ALGORITMA *MULTINOMIAL NAÏVE BAYES*

Disusun untuk Melengkapi Tugas dan Memenuhi Syarat Kelulusan
Program Strata 1 Jurusan Teknik Informatika Fakultas Teknik
Universitas Muhammadiyah Jember



Robbi Zulcharnain

2010651076

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

2025

HALAMAN PERSETUJUAN SIDANG TUGAS AKHIR

SENTIMEN ANALISIS ULASAN DUOLINGO DENGAN METODE ALGORITMA *MULTINOMIAL NAÏVE BAYES*

Oleh

Robbi Zulcharnain

2010651076

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



Ginanjjar Abdurrahman, S.Si., M.Pd
NIDN. 0714078704

Daryanto, S.Kom., M.Kom
NIDN. 0707077203

SURAT PERNYATAAN

Yang bertanda tangan di bawah ini:

Nama : Robbi Zulcharnain

NIM : 2010651076

Menyatakan dengan sesungguhnya bahwa Tugas Akhir berjudul "**SENTIMEN ANALISIS ULASAN DUOLINGO DENGAN METODE ALGORITMA MULTINOMIAL NAÏVE BAYES**" adalah benar merupakan karya sendiri. Hal-hal yang bukan karya saya, dalam skripsi tersebut diberi tanda citasi dan ditunjukkan dalam daftar pustaka.

Apabila di kemudian hari terbukti pernyataan saya tidak benar dan ditemukan pelanggaran atas karya Tugas Akhir ini, saya bersedia menerima sanksi akademik berupa pencabutan Tugas Akhir dan gelar saya peroleh dari Tugas Akhir tersebut.

Jember, 23 Januari 2025

Yang membuat pernyataan



Robbi Zulcharnain
NIM. 2010651076

LEMBAR PENGESAHAN

SENTIMEN ANALISIS ULASAN DUOLINGO DENGAN METODE
ALGORITMA *MULTINOMIAL NAÏVE BAYES*

Oleh

Robbi Zulcharnain

2010651076

Telah mempertanggung jawabkan Laporan Tugas Akhirnya pada sidang Tugas Akhir Tanggal 18 Januari 2025 sebagai salah satu syarat kelulusan dan mendapatkan gelar Sarjana Komputer (S.Kom)

Disetujui oleh,

Dosen Penguji:
Penguji I



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

Dosen Pembimbing:
Pembimbing I



Ginanjar Abdurrahman, S.Si., M.Pd
NIDN. 0714078704

Penguji II



Habibatul Azizah Al Faruq, M.Pd
NIDN. 0718128901

Pembimbing II



Darvanto, S.Kom., M.Kom
NIDN. 0707077203

Mengesahkan,
Dekan
Fakultas Teknik



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

Mengetahui,
Ketua Program Studi
Teknik Informatika



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

MOTTO

“Setiap detik adalah kesempatan untuk mengubah hidupmu.”

-Robbi Zulcharnain

"Setiap langkah kecil yang diambil hari ini adalah fondasi besar untuk impian di masa depan."

-Robbi Zulcharnain

"Entah seekor ikan hidup di sungai yang jernih atau di selokan, selama ia terus berenang maju, ia akan tumbuh dengan indah."

-Koro-Sensei, Assassination Classroom

"Satu-satunya cara untuk melakukan pekerjaan hebat adalah dengan mencintai apa yang Anda lakukan."

-Steve Jobs

"Jika Anda tidak bisa melakukan hal-hal besar, lakukan hal kecil dengan cara yang hebat."

-Napoleon Hill

"A lesson in fighting is not simply about fighting others; it's about fighting yourself."

-Roronoa Zoro, One Piece

PERSEMBAHAN

Dengan segala rasa syukur kepada Allah SWT atas limpahan rahmat, hidayah, dan kasih sayang-Nya, sehingga saya dapat menyelesaikan Tugas Akhir yang berjudul:

“SENTIMEN ANALISIS ULASAN DUOLINGO DENGAN METODE ALGORITMA *MULTINOMIAL NAÏVE BAYES*”

Penyusunan Tugas Akhir ini tidak lepas dari dukungan, bantuan, dan saran. Oleh karena itu, saya persembahkan dengan penuh cinta dan rasa hormat kepada:

1. Ayah dan Ibu tercinta saya M. Romli dan Sumiati, yang selalu menjadi sumber kekuatan, inspirasi, dan doa tanpa henti. Perjuangan ini adalah bentuk kecil dari rasa terima kasih atas semua pengorbanan dan kasih sayang.
2. Kakak-kakak tercinta saya Eka Rahayu, Dewi Indriani dan Sherly Tri Purnamasari, yang selalu memberikan dukungan dan semangat di setiap langkah perjalanan akademik saya.
3. Dosen Pembimbing saya Bapak Ginanjar Abdurrahman dan Bapak Daryanto, yang dengan penuh kesabaran, bimbingan, dan arahan membantu saya menyelesaikan penelitian ini.
4. Rekan-rekan seperjuangan saya Bramestya, Alan, Albert, Dinar, Mirza, dan Thezar yang selalu hadir memberikan motivasi, semangat, dan kebersamaan dalam menghadapi berbagai tantangan selama masa kuliah.
5. Partner saya Dwi Ufaira, seseorang yang selalu memberikan dukungan dan selalu menghibur saya (walaupun sering menghilang). Meskipun sering bercanda tetapi tetap memberikan dukungan kepada penulis.

Semoga karya ini dapat menjadi manfaat bagi perkembangan ilmu pengetahuan, khususnya dalam bidang teknologi dan analisis data.

KATA PENGANTAR

Puji dan syukur kepada Allah SWT atas segala rahmat, dan karunianya sehingga penulis dapat menyelesaikan skripsi ini yang berjudul "**Sentimen Analisis Ulasan Duolingo dengan Metode Algoritma *Multinomial Naïve Bayes***." Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer (S.Kom) di Program Studi Teknik Informatika Fakultas Teknik Universitas Muhammadiyah Jember.

Penyusunan skripsi ini tidak terlepas dari dukungan, bantuan, dan bimbingan dari berbagai pihak. Oleh karena itu, penulis ingin menyampaikan rasa terima kasih yang sebesar-besarnya kepada:

1. Rektor Universitas Muhammadiyah Jember
2. Dekan Fakultas Teknik Universitas Muhammadiyah Jember.
3. Ibu Rosita Yanuarti, S.Kom., M. Cs, selaku ketua Program Studi Teknik Informatika.
4. Bapak Ginanjar Abdurrahman, S.Si., M.Pd, dan Bapak Daryanto, S.Kom., M.Kom. selaku dosen pembimbing 1 dan dosen pembimbing 2.
5. Bapak Hardian Oktavianto, S.Si., M.Kom, dan Ibu Habibatul Azizah Al Faruq, M.Pd, selaku dosen penguji 1 dan dosen penguji 2.
6. Semua pihak yang tidak dapat disebutkan satu per satu, yang telah membantu penulis baik secara langsung maupun tidak langsung.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Oleh karena itu, penulis sangat mengharapkan kritik dan saran yang membangun untuk penyempurnaan karya ini. Semoga skripsi ini dapat memberikan manfaat bagi pembaca dan kontribusi bagi perkembangan ilmu pengetahuan.

Akhir kata, penulis mengucapkan terima kasih atas perhatian dan apresiasi yang diberikan. Semoga karya ini menjadi langkah awal bagi pencapaian yang lebih besar di masa depan.

Jember, Januari 2025



Robbi Zulcharnain
2010651076

DAFTAR ISI

HALAMAN PERSETUJUAN SIDANG TUGAS AKHIR.....	iii
SURAT PERNYATAAN.....	iv
LEMBAR PENGESAHAN	v
MOTTO	vi
PERSEMBAHAN.....	vii
ABSTRAK	viii
ABSTRACT	ix
KATA PENGANTAR.....	x
DAFTAR ISI	xi
DAFTAR GAMBAR	xiv
DAFTAR TABEL	xvi
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	5
BAB II TINJAUAN PUSTAKA.....	6
2.1 Penelitian Terdahulu.....	6
2.2 Kajian Teori.....	9
2.2.1 <i>Google Play Store</i>	9
2.2.2 <i>Duolingo</i>	9
2.2.3 <i>Machine Learning</i>	10
2.2.4 <i>Text Mining</i>	11
2.2.5 <i>Text Preprocessing</i>	11
2.2.2.1 <i>Cleansing</i>	12
2.2.2.2 <i>Case Folding</i>	12
2.2.2.3 <i>Tokenizing</i>	12
2.2.2.4 <i>Stemming</i>	13

2.2.2.5	<i>Stopword</i>	13
2.2.6	Analisis Sentimen	13
2.2.7	VADER (<i>Valanced Aware Dictionary Sentiment Reasoner</i>).....	14
2.2.8	<i>Naïve Bayes</i>	15
2.2.9	<i>Multinomial Naïve Bayes</i>	16
2.2.10	<i>K-Fold Cross Validation</i>	19
2.2.11	Ketepatan Klasifikasi	20
2.2.12	Visualisasi <i>Word Cloud</i>	21
BAB III	METODOLOGI PENELITIAN	24
3.1	Kerangka Penelitian.....	24
3.2	Alur Penelitian.....	25
3.3	<i>Crawling</i>	25
3.4	<i>Preprocessing</i>	25
3.5	<i>Multinomial Naïve Bayes</i>	28
3.6	<i>K-Fold Cross Validation</i>	33
3.7	Visualisasi <i>Word Cloud</i>	35
3.8	<i>Predicted Sentiment</i>	35
BAB IV	HASIL DAN PEMBAHASAN.....	36
4.1	Pengambilan Data.....	36
4.2	<i>Contextualized Representation</i>	38
4.3	<i>Preprocessing</i>	38
4.4	<i>Vectorizer</i>	42
4.5	Implementasi Klasifikasi <i>Multinomial Naïve Bayes</i>	43
4.6	<i>K-Fold Cross Validation</i>	48
4.7	<i>Word Cloud</i>	55
4.8	Uji Model.....	57
BAB V	KESIMPULAN DAN SARAN	60
5.1	Kesimpulan.....	60
5.2	Saran.....	60
DAFTAR PUSTAKA		61

LAMPIRAN..... 66
BIODATA PENULIS..... 68



DAFTAR GAMBAR

Gambar 1. 1 Beberapa ulasan dari pengguna pada Duolingo	2
Gambar 2. 1 Visualisasi <i>Word Cloud</i>	22
Gambar 3. 1 Alur Metode Penelitian.....	24
Gambar 3. 2 <i>Labelling</i>	26
Gambar 3. 3 Seleksi Fitur.....	26
Gambar 4. 1 Proses <i>install library google play scraper</i>	36
Gambar 4. 2 Proses ekstrak data	36
Gambar 4. 3 Proses menyimpan data	37
Gambar 4. 4 Proses install <i>library VADER</i>	37
Gambar 4. 5 Proses penentuan sentimen.....	37
Gambar 4. 6 Alur penelitian	38
Gambar 4. 7 Proses <i>labelling</i>	39
Gambar 4. 8 Proses seleksi fitur.....	39
Gambar 4. 9 Proses <i>Cleansing</i>	40
Gambar 4. 10 Proses <i>Case Folding</i>	40
Gambar 4. 11 Proses <i>Tokenizing</i>	41
Gambar 4. 12 Proses <i>Stopword</i>	41
Gambar 4. 13 Proses <i>Stemming</i>	42
Gambar 4. 14 Hasil <i>preprocessing</i>	42
Gambar 4. 15 Proses memisahkan data.....	43
Gambar 4. 16 Visualisasi <i>vectorizer heatmap</i>	43
Gambar 4. 17 Hasil model <i>Multinomial Naïve Bayes</i>	43
Gambar 4. 18 Menentukan <i>prior</i>	44
Gambar 4. 19 Hasil frekuensi kata positif dan negatif serta total kata unik	44
Gambar 4. 20 Hasil perhitungan <i>likelihood</i> perkata.....	45
Gambar 4. 21 Hasil perhitungan <i>posterior probability</i>	45
Gambar 4. 22 Hasil perhitungan <i>Cmap</i>	46
Gambar 4. 23 Frekuensi kata dalam kategori positif dan negatif	47
Gambar 4. 24 Teknik <i>add-one</i> atau <i>laplace smoothing</i> pada semua dokumen	48
Gambar 4. 25 Mengonversi kedalam bentuk <i>string</i>	49
Gambar 4. 26 <i>K-Fold Cross Validation setup</i>	49
Gambar 4. 27 Evaluasi tahap <i>K-Fold Cross Validation</i>	49
Gambar 4. 28 Menghitung <i>accuracy</i> dan standar deviasi	49
Gambar 4. 29 Hasil <i>accuracy</i> setiap fold	50
Gambar 4. 30 Hasil klasifikasi skenario 1	51
Gambar 4. 31 Hasil klasifikasi skenario 2	52
Gambar 4. 32 Hasil klasifikasi skenario 3	53
Gambar 4. 33 Hasil klasifikasi skenario 4	54

Gambar 4. 34 Hasil klasifikasi skenario 5	55
Gambar 4. 35 Hasil <i>word cloud</i> frekuensi kata negatif.....	56
Gambar 4. 36 Hasil <i>word cloud</i> frekuensi kata positif.....	56
Gambar 4. 37 Proses <i>Confusion Matrix</i>	57
Gambar 4. 38 Menampilkan hasil <i>Confusion Matrix</i>	57
Gambar 4. 39 Hasil <i>Confusion Matrix</i>	58
Gambar 4. 40 Prediksi sentimen positif	59
Gambar 4. 41 Prediksi sentimen negatif.....	59



DAFTAR TABEL

Tabel 2. 1 Tabel proses <i>Cleansing</i>	12
Tabel 2. 2 Tabel Proses <i>Case Folding</i>	12
Tabel 2. 3 Tabel Proses <i>Tokenizing</i>	12
Tabel 2. 4 Tabel Proses <i>Stemming</i>	13
Tabel 2. 5 Proses <i>Stopword</i>	13
Tabel 2. 6 Ilustrasi dalam pembagian data.....	20
Tabel 2. 7 Proses <i>Confusion Matrix</i>	20
Tabel 3. 1 Proses <i>Cleansing</i>	27
Tabel 3. 2 Proses <i>Case Folding</i>	27
Tabel 3. 3 Proses <i>Tokenizing</i>	27
Tabel 3. 4 Proses <i>Stopword</i>	28
Tabel 3. 5 Proses <i>Stemming</i>	28
Tabel 3. 6 Hasil perhitungan <i>likelihood</i>	29
Tabel 3. 7 Hasil perhitungan <i>Cmap</i>	32
Tabel 3. 8 Pembagian data uji dan data latih.....	34
Tabel 3. 9 <i>Confusion Matrix</i> data uji 20%	34
Tabel 4. 1 Hasil perhitungan total	46
Tabel 4. 2 <i>Classification report</i> skenario 1	50
Tabel 4. 3 <i>Classification report</i> skenario 2	51
Tabel 4. 4 <i>Classification report</i> skenario 3	52
Tabel 4. 5 <i>Classification report</i> skenario 4	53
Tabel 4. 6 <i>Classification report</i> skenario 5	54

LAMPIRAN

Lampiran 1 <i>Source Code</i> Penelitian.....	66
Lampiran 2 Surat Pernyataan Publikasi.....	67

