

**IMPLEMENTASI ALGORITMA NAIVE BAYES UNTUK ANALISIS
SENTIMEN REVIEW PRODUK DI TOKOPEDIA**



Diajukan sebagai salah satu syarat untuk memperoleh
Gelar Sarjana Komputer (S.Kom.)

Rezkhi Alfianda

0000011483

UMMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA**

TANGERANG

2023

**IMPLEMENTASI ALGORITMA NAIVE BAYES UNTUK ANALISIS
SENTIMEN REVIEW PRODUK DI TOKOPEDIA**



Diajukan sebagai salah satu syarat untuk memperoleh
Gelar Sarjana Komputer (S.Kom.)

Rezki Alifianda
00000011483

UMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA

PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA

TANGERANG

2023

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Rezkhi Alifianda
Nomor Induk Mahasiswa : 00000011483
Program Studi : Informatika

Skripsi dengan judul:

Implementasi Algoritma Naive Bayes Untuk Analisis Sentimen Review Produk di Tokopedia

merupakan hasil karya saya sendiri bukan plagiat dari karya ilmiah yang ditulis oleh orang lain, dan semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar serta dicantumkan di Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/ penyimpangan, baik dalam pelaksanaan Skripsi maupun dalam penulisan laporan Skripsi, saya bersedia menerima konsekuensi dinyatakan TIDAK LULUS untuk Tugas akhir yang telah saya tempuh.

Tangerang, 3 Januari 2023



(Rezkhi Alifianda)

UMM
UNIVERSITAS
MULTIMEDIA
NUSANTARA

HALAMAN PENGESAHAN

Skripsi dengan judul

Implementasi Algoritma Naive Bayes Untuk Analisis Sentimen Review Produk di
Tokopedia

oleh

Nama : Rezkhi Alifianda
NIM : 00000011483
Program Studi : Informatika
Fakultas : Teknik dan Informatika

Telah diujikan pada hari Selasa, 10 Januari 2023

Pukul 13.00 s/d 15.00 dan dinyatakan

LULUS

Dengan susunan penguji sebagai berikut

Ketua Sidang



Digitally signed
by Wirawan
Date: 2023.01.19
08:32:14 +07'00'

(Wirawan Istiono, S.Kom., M.Kom.)

NIDN: 0313048304

Penguji



Digitally signed
by Fenina Adline
Twince Tobing

(Fenina Adline Twince Tobing, S.Kom.,

M.Kom)

NIDN: 0406058802

Pembimbing



Digitally signed by Ni
Made Satvika Iswari
Date: 2023.01.20
08:12:58 +08'00'

(Dr. Ni Made Satvika Iswari, S.T., M.T.)

NIDN: 0306019001

Ketua Program Studi Informatika,



Digitally signed
by Marlinda
Vasty Overbeek
Date: 2023.01.20
10:01:53 +07'00'

(Marlinda Vasty Overbeek, S.Kom., M.Kom.)

NIDN: 0818038501

iii

**HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK
KEPENTINGAN AKADEMIS**

Sebagai sivitas akademik Universitas Multimedia Nusantara, saya yang bertanda tangan di bawah ini:

Nama : Rezghi Alifianda
NIM : 00000011483
Program Studi : Informatika
Fakultas : Teknik dan Informatika
Jenis Karya : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada **Universitas Multimedia Nusantara** hak Bebas Royalti Non-eksklusif (*Non-exclusive Royalty-Free Right*) atas karya ilmiah saya yang berjudul:

**IMPLEMENTASI ALGORITMA NAIVE BAYES UNTUK ANALISIS
SENTIMEN REVIEW PRODUK DI TOKOPEDIA**

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Non eksklusif ini Universitas Multimedia Nusantara berhak menyimpan, mengalih media / format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis / pencipta dan sebagai pemilik Hak Cipta. Demikian pernyataan ini saya buat dengan sebenarnya.

Tangerang, 3 Januari 2023
Yang menyatakan



Rezghi Alifianda

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

Halaman Persembahan / Motto

”Berpikirlah positif, tidak peduli seberapa keras kehidupanmu.”

Ali bin Abi Thalib



KATA PENGANTAR

Puji Syukur atas berkat dan rahmat kepada Tuhan Yang Maha Esa, atas selesainya penulisan laporan Skripsi ini dengan judul: Implementasi Algoritma Naive Bayes Untuk Analisis Sentimen Review Produk di Tokopedia dilakukan untuk memenuhi salah satu syarat untuk mencapai gelar Sarjana/Magister Komputer Jurusan Informatika Pada Fakultas Teknik dan Informatika Universitas Multimedia Nusantara. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan skripsi ini, sangatlah sulit bagi saya untuk menyelesaikan skripsi ini. Oleh karena itu, saya mengucapkan terima kasih kepada:

1. Bapak Dr. Ninok Leksono, selaku Rektor Universitas Multimedia Nusantara.
2. Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Dekan Fakultas Teknik dan Informatika Universitas Multimedia Nusantara.
3. Ibu Marlinda Vasty Overbeek, S.Kom., M.Kom., selaku Ketua Program Studi Informatika Universitas Multimedia Nusantara.
4. Ibu Dr. Ni Made Satvika Iswari, S.T., M.T., sebagai Pembimbing yang telah banyak meluangkan waktu untuk memberikan bimbingan, arahan dan motivasi atas terselesainya tesis ini.
5. Orang Tua, adik, dan teman saya yang telah memberikan bantuan dukungan material dan moral, sehingga penulis dapat menyelesaikan tesis ini.

Semoga skripsi ini bermanfaat, baik sebagai sumber informasi maupun sumber inspirasi, bagi para pembaca.

Tangerang, 3 Januari 2023



Rezki Alifianda

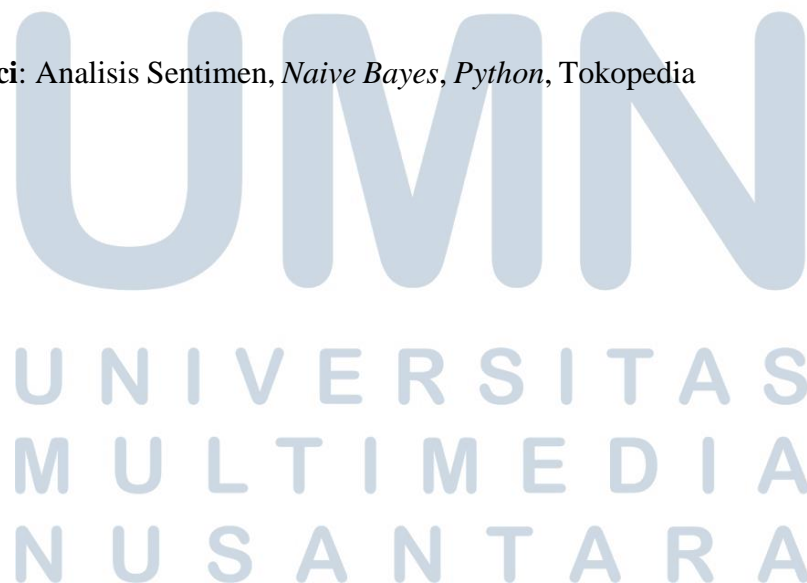
IMPLEMENTASI ALGORITMA NAIVE BAYES UNTUK ANALISIS SENTIMEN REVIEW PRODUK DI TOKOPEDIA

Rezki Alifianda

ABSTRAK

Latar Belakang Tokopedia merupakan sebuah platform jual beli berbasis online yang memungkinkan setiap individu dan pemilik bisnis di Indonesia untuk mengembangkan dan memasarkan produk yang ingin dijual dengan mudah dan aman. Hasil review dari produk yang sudah dibeli oleh *user* dapat dikategorikan dengan review positif, negatif, dan netral. Review atau ulasan-ulasan dari produk atau barang yang ada pada Tokopedia jumlahnya sangat banyak, ribuan bahkan jutaan review atau ulasan yang ada pada Tokopedia. Sehingga untuk melakukan klasifikasi pada review atau ulasan secara manual akan memakan banyak waktu. Oleh karena itu, diperlukan program untuk melakukan klasifikasi pada review atau ulasan-ulasan dengan menggunakan analisis sentimen. Model klasifikasi menggunakan algoritma Naive Bayes. Penelitian ini bertujuan untuk mengimplementasikan algoritma Naive Bayes dengan menggunakan pemrograman Python serta menghitung tingkat akurasi prediksi dari sistem yang dibuat. Berdasarkan penelitian sebelumnya, algoritma Naive Bayes memberikan performa paling baik bila dibandingkan dengan algoritma klasifikasi lainnya. Hasil pengujian dengan performa terbaik dengan perbandingan *data train* sebesar 70% dan *data test* sebesar 30% dengan nilai *accuracy* 83%, *precision* 85%, *recall* 91%, dan *f1-score* 88%.

Kata kunci: Analisis Sentimen, *Naive Bayes*, *Python*, Tokopedia



Implementation of the Naive Bayes Algorithm for Product Review Sentiment Analysis at Tokopedia

Rezkhi Alifianda

ABSTRACT

Background research Tokopedia is an online-based buying and selling platform that allows every individual and business owner in Indonesia to develop and market the products they want to sell easily and safely. The review results of products that have been purchased by users can be categorized as positive, negative, or neutral reviews. There are very many reviews or reviews of products or goods on Tokopedia—thousands or even millions of reviews or reviews on Tokopedia. So to classify reviews or reviews manually will take a lot of time. Therefore, a program is needed to classify reviews using sentiment analysis. The classification model uses the Naive Bayes algorithm. This study aims to implement the Naive Bayes algorithm using Python programming and calculate the prediction accuracy level of the system created. Based on previous research, the Naive Bayes algorithm provides the best performance when compared to other classification algorithms. The test results with the best performance with a comparison of data train of 70% and data test of 30% with a value of accuracy 83%, precision 85%, recall 91%, and f1-score 88%.

Keywords: Naive Bayes, Python, Sentiment Analysis, Tokopedia



DAFTAR ISI

HALAMAN JUDUL	i
PERNYATAAN TIDAK MELAKUKAN PLAGIAT	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERSETUJUAN PUBLIKASI ILMIAH	iv
HALAMAN PERSEMBAHAN/MOTO	v
KATA PENGANTAR	vi
ABSTRAK.....	vii
ABSTRACT.....	viii
DAFTAR ISI.....	ix
DAFTAR GAMBAR	x
DAFTAR TABEL.....	xi
DAFTAR LAMPIRAN.....	xii
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	2
1.3 Batasan Permasalahan	3
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.6 Sistematika Penulisan	3
BAB 2 LANDASAN TEORI	5
2.1 Tokopedia	5
2.2 Analisis Sentimen	5
2.3 <i>Text Preprocessing</i>	5
2.3.1 <i>Case Folding</i>	5
2.3.2 <i>Tokenizing</i>	5
2.3.3 <i>Filtering</i>	6
2.3.4 <i>Stemming</i>	6
2.4 <i>Term Frequency-Inverse Document Frequency</i>	6
2.5 <i>Confusion Matrix</i>	6
2.6 Algoritma Naive Bayes	7
BAB 3 METODOLOGI DAN PERANCANGAN SISTEM	9
3.1 Metode Penelitian	9
3.2 Perancangan Sistem	9
3.2.1 <i>Flowchart</i> Utama	10
3.2.2 <i>Flowchart Preprocessing</i>	11
3.2.3 <i>Flowchart Model Klasifikasi Naive Bayes</i>	13
BAB 4 HASIL DAN DISKUSI	14
4.1 Spesifikasi Sistem	14
4.2 Implementasi Sistem	15
4.3 Uji Coba	27
4.3.1 Skenario Uji Coba	27
4.3.2 Hasil Uji Coba	28
4.3.3 Evaluasi Hasil Uji Coba	28
BAB 5 SIMPULAN DAN SARAN.....	31
1.1 Kesimpulan	31
1.2 Saran	31
DAFTAR PUSTAKA	32

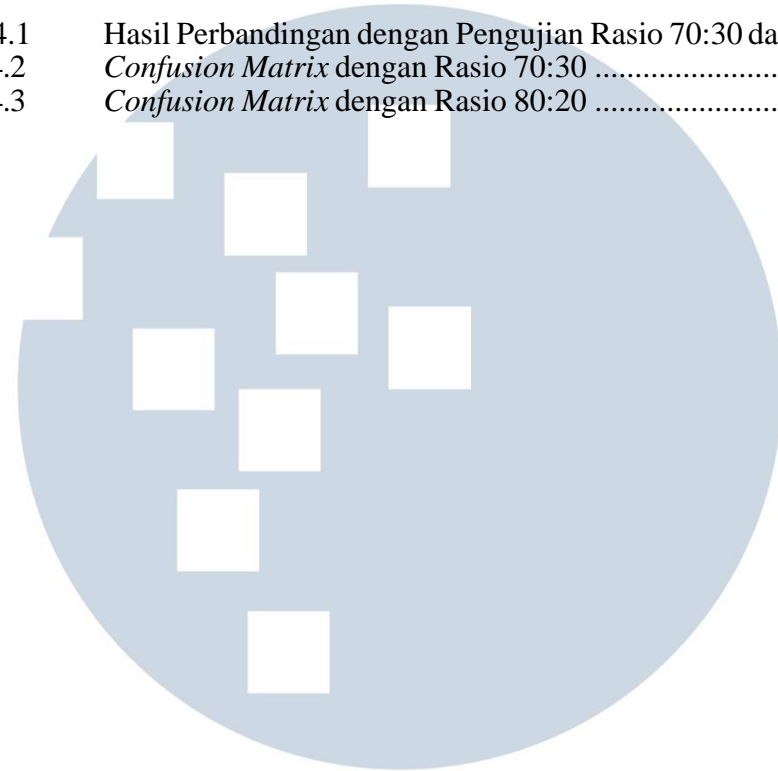
DAFTAR GAMBAR

Gambar 1.1	Data Pengunjung Website E-Commerce Bulanan Terbanyak	1
Gambar 3.1	Flowchart Utama	11
Gambar 3.2	Flowchart Preprocessing	12
Gambar 3.3	Flowchart Naive Bayes	13
Gambar 4.1	Potongan Kode Import Library	15
Gambar 4.2	Potongan Kode Import Dataset	15
Gambar 4.3	Dataset	16
Gambar 4.4	Simbol dan Karakter Tanda Baca	16
Gambar 4.5	Potongan Kode Cleaning Text	16
Gambar 4.6	Potongan Kode <i>Case Folding</i>	17
Gambar 4.7	Hasil <i>Case Folding</i>	17
Gambar 4.8	Potongan Kode <i>Tokenization</i>	18
Gambar 4.9	Hasil <i>Tokenization</i>	18
Gambar 4.10	Potongan Kode <i>Stopwords</i>	19
Gambar 4.11	Hasil <i>stopwords</i>	19
Gambar 4.12	Potongan Kode <i>Stemming</i>	20
Gambar 4.13	Hasil <i>Stemming</i>	20
Gambar 4.14	Potongan Kode Labeling dari Rating	21
Gambar 4.15	Hasil Labeling dari Rating	21
Gambar 4.16	Jumlah dari Hasil Positif, Negatif, dan Netral	22
Gambar 4.17	Potongan Kode Labeling dengan Text Blob	22
Gambar 4.18	Potongan Kode <i>getSubjectivity</i>	23
Gambar 4.19	Potongan Kode <i>getPolarity</i>	23
Gambar 4.20	Potongan Kode Membuat Kolom Subjectivity dan Polarity	23
Gambar 4.21	Hasil dari Subjectivity dan Polarity	24
Gambar 4.22	Potongan Kode <i>getAnalysis</i> untuk menentukan hasil sentimen	24
Gambar 4.23	Hasil Analisis	25
Gambar 4.24	Klasifikasi Positif	26
Gambar 4.25	Klasifikasi Netral	26
Gambar 4.26	Klasifikasi Negatif	27

U N I V E R S I T A S
 M U L T I M E D I A
 N U S A N T A R A

DAFTAR TABEL

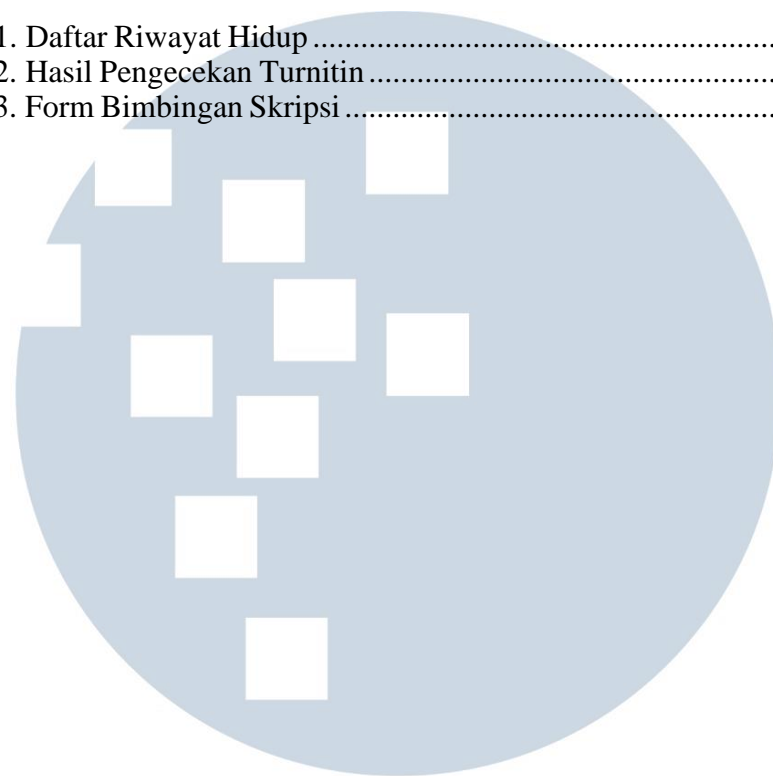
Tabel 4.1	Hasil Perbandingan dengan Pengujian Rasio 70:30 dan 80:20	28
Tabel 4.2	<i>Confusion Matrix</i> dengan Rasio 70:30	28
Tabel 4.3	<i>Confusion Matrix</i> dengan Rasio 80:20	28



UMMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR LAMPIRAN

Lampiran 1. Daftar Riwayat Hidup	33
Lampiran 2. Hasil Pengecekan Turnitin	35
Lampiran 3. Form Bimbingan Skripsi	35



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA