

**ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI ZALORA
DI GOOGLE PLAY STORE MENGGUNAKAN ALGORITMA
NAIVE BAYES**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

TEDDY MUKHASAN

20.11.3668

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2024

**ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI
ZALORA DI GOOGLE PLAY STORE MENGGUNAKAN
ALGORITMA NAIVE BAYES**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

TEDDY MUKHASAN

20.11.3668

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2024

HALAMAN PERSETUJUAN

SKRIPSI

**ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI ZALORA DI
GOOGLE PLAY STORE MENGGUNAKAN ALGORITMA NAIVE BAYES**


yang disusun dan diajukan oleh

Teddy Mukhasan

20.11.3668

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal < 19 Juli 2024 >

Dosen Pembimbing,


Ike Verawati, M.Kom

NIK. 190302237

HALAMAN PENGESAHAN

SKRIPSI

ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI ZALORA DI
GOOGLE PLAY STORE MENGGUNAKAN ALGORITMA NAIVE BAYES

yang disusun dan diajukan oleh

Teddy Mukhsan

20.11.3668

Telah dipertahankan di depan Dewan Penguji
pada tanggal <19 Juli 2024>

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Robert Marco, M.T., Ph.D
NIK. 190302228

Rizqi Sukma Khurisma, M.Kom
NIK. 190302215

Ihs Yerawati, M.Kom
NIK. 190302237

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal < 19 Juli 2024 >

DEKAN FAKULTAS ILMU KOMPUTER



Hanif Al Fatah, M.Kom., Ph.D.
NIK. 190302096

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : TEDDY MUKHASAN
NIM : 20.11.3668

Menyatakan bahwa Skripsi dengan judul berikut:

**ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI ZALORA DI
GOOGLE PLAY STORE MENGGUNAKAN ALGORITMA NAIVE BAYES**

Dosen Pembimbing : Ike Verawati, M.Kom

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari 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 tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian 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.

Yogyakarta, < 19 Juli 2024 >

Yang Menyatakan,



TEDDY MUKHASAN

Teddy Mukhasan

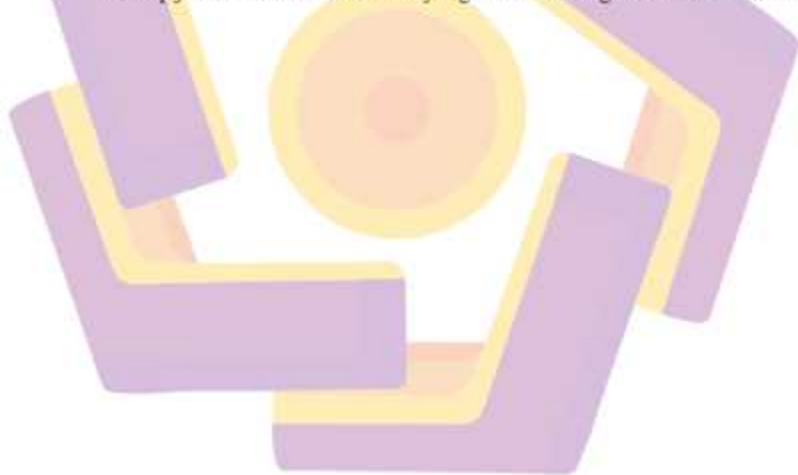
HALAMAN PERSEMBAHAN

Alhamdulillah, Segala Puji Bagi ALLAH SWT. Kepada-Nya kita memohon, memuji dan meminta perlindungan, serta petunjuk Kepada-Nya. Kita berlindung kepada ALLAH SWT dari kejahatan diri kita dan keburukan amal-amal kita. Dengan ini penulis mempersembahkan skripsi ini kepada semua pihak yang telah membantu dan memberikan motivasi serta dukungan yang besar sehingga skripsi ini dapat diselesaikan dengan sebaik baiknya.

Melalui skripsi ini saya persembahkan terimakasih serta syukur kepada:

1. Rasa syukur kepada Allah SWT yang telah memberikan rahmat serta karunia untuk terus mengucap syukur dan sabar dalam menyelesaikan tugas akhir skripsi.
2. Kepada keluarga penulis Bapak Hairawan, Ibu Sri Hastuti, Bapak Agus Supratman dan nenek dari penuli Ibu Murdiah yang selalu memberikan doa, cinta dan dukungan yang tidak pernah berhenti diberikan kepada penulis.
3. Ibu Ike Verawati, M.Kom selaku dosen pembimbing yang selalu memberikan saran dan masukan kepad penulis dalam menyelesaikan skripsi ini.
4. Bapak/Ibu dosen yang telah memberikan ilmu dan pendidikan selama berkuliah di Amikom.
5. Terima kasih kepada Isratul Aini Azhari yang selalu kebersamai penulis sampai dengan tugas akhir skripsi ini. Walaupun dalam proses penulisan skripsi ini banyak hal yang terlewat dan harus dikorbankan. Semoga kedepannya banyak hal baik dan menyenangkan yang harus bisa dilewati bersama.
6. Staf UPT LAB dan teman teman SS UPT Lab yang selalu memberikan support dalam mencapai kesuksesan.
7. Teman teman HMIF yang menjadi pemberi dukungan dalam mencapai kesuksesan yang penuh suka dan duka. Terima kasih banyak telah membantu banyak hal yang tidak diketahui sampai dengan sekarang penulis bisa menyelesaikan tugas akhir skripsi ini.

8. Teman teman Senat Mahasiswa periode 2021/2022 yang selalu mengingatkan penulis terkait skripsi.
9. Kepada teman teman “SUBEDU” yang hampir setiap saat memberikan segala yang mereka bisa kepada penulis sehingga dapat menyelesaikan tugas akhir skripsi.
10. Yang terakhir terima kasih kepada laki-laki yang selalu memiliki pemikiran akan hal hal kecil yang sulit dimengerti, sang penulis sebuah karya tulis ini, diri saya sendiri, Teddy Mukhasan. Seorang laki-laki yang berusia 22 tahun saat menyelesaikan tugas akhir skripsi. Terima kasih kepada diri saya sendiri yang tetap mengupayakan semua hal dan terus berusaha memperbaiki segala hal terkait tugas akhir skripsi ini. “Jika memiliki tujuan pasti ada satu dua hal yang terlewat, fokus saja apa yang ada di depan. Cukup jadikan masa lalu atau hal yang terlewat sebagai bahan evaluasi diri”.



KATA PENGANTAR

Puji syukur kehadirat ALLAH SWT karena berkat rahmat serta hidayah-Nya penulis diberikan kesehatan dan kekuatan serta iman yang kuat sehingga dapat menyelesaikan skripsi yang berjudul **“ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI ZALORA DI GOOGLE PLAY MENGGUNAKAN ALGORITMA NAÏVE BAYES”**. Oleh karena itu penulis mengucapkan syukur karena dapat menyelesaikan skripsi ini. Skripsi ini diajukan kepada Program Studi S1 Informatika, Fakultas Ilmu Komputer, serta Universitas AMIKOM Yogyakarta.

Penulis dalam menyelesaikan skripsi mendapatkan beberapa hambatan dalam berbagai hal, namun banyak pihak yang selalu membantu sehingga skripsi ini dapat diselesaikan dengan baik. Dengan begitu penulis hendak mengungkapkan terima kasih terhadap:

1. Bapak prof. Dr. M Suyanto, M.M selaku Rektor Universitas AMIKOM Yogyakarta.
2. Bapak Hanif Al Fatta, S.Kom., M.Kom selaku Dekan Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta.
3. Ibu Windha Mega Pradya D, M.Kom., selaku Ketua Program Studi S1 Informatika Universitas AMIKOM Yogyakarta.
4. Ibu Ike Verawati, M.Kom selaku dosen pembimbing yang selalu memberikan saran dan arahan dalam proses penulisan skripsi ini.
5. Bapak Hairawan dan Ibu Sri Hastuti selaku orang tua serta Ibu Murdiah selaku wali dengan Bapak Agus Supratman yang selalu memberikan doa dan motivasi kepada penulis.

Yogyakarta, < 18 September 2024 >

Penulis

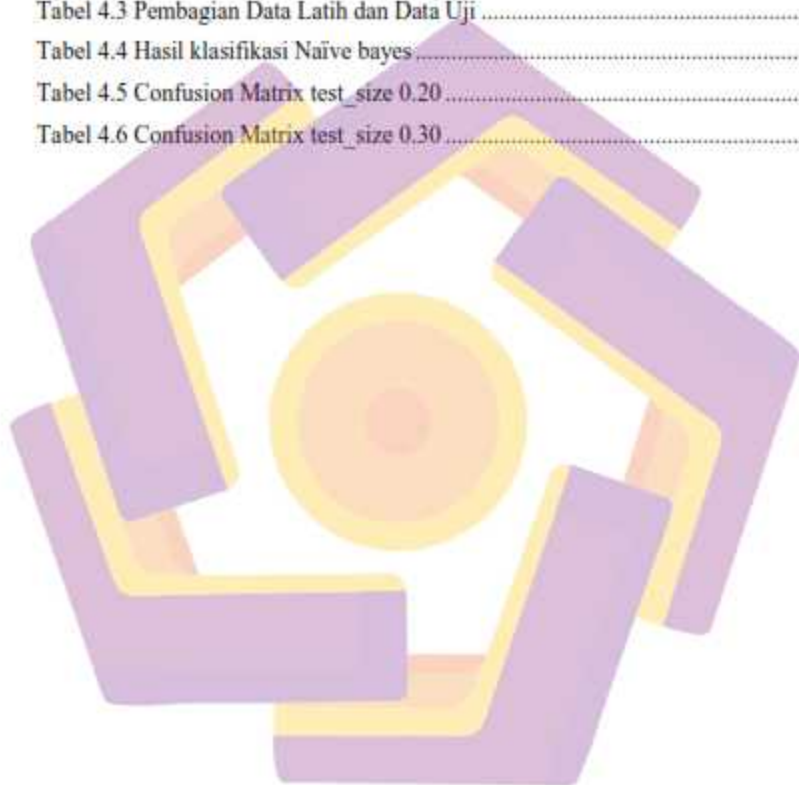
DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vii
DAFTAR ISI.....	viii
DAFTAR TABEL.....	x
DAFTAR GAMBAR	xi
INTISARI	xii
<i>ABSTRACT</i>	xiii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah.....	1
1.3 Batasan Masalah	2
1.4 Tujuan Penelitian.....	2
1.5 Manfaat Penelitian	2
1.6 Sistematika Penulisan	3
BAB II TINJAUAN PUSTAKA	4
2.1 Studi Literatur.....	4
2.2 Dasar Teori	12
2.2.1. Zalora	12
2.2.2. Naïve Bayes	12
2.2.3. Scraping	13
2.2.4. Analisis Sentimen	13
2.2.5. Python	14
2.2.6. Klasifikasi	14
2.2.7. Pre-Processing.....	14
2.2.8. Pembobotan TF-IDF	15

2.2.9. Pelabelan	15
2.2.10. Confusion Matrix	15
BAB III METODE PENELITIAN	17
3.1 Alur Penelitian	17
3.2 Alat dan Bahan.....	19
BAB IV HASIL DAN PEMBAHASAN	21
4.1. Pengambilan Dataset	21
4.2. Pelabelan Data	23
4.3. Data Cleaning	24
4.4. Pre-Processing	25
4.4.1. Case Folding	25
4.4.2. Stopword Removal	26
4.4.3. Tokenizing	27
4.4.4. Stemming	28
4.5. Pembobotan TF-IDF	29
4.6. <i>Training & Testing</i> Dataset	30
4.7. Klasifikasi Algoritma Naive Bayes	31
4.8. Evaluasi	32
4.8.1. Confusion matrix test_size 0.20	32
4.8.2. Confusion matrix test_size 0.30	34
BAB V PENUTUP	37
5.1 Kesimpulan	37
5.2 Saran	37
REFERENSI	39

DAFTAR TABEL

Tabel 2.1 Keaslian Penelitian.....	6
Tabel 3.1 Deskripsi Variabel	20
Tabel 4. 1 dataset	21
Tabel 4.2 Dataset Ulasan Review	23
Tabel 4.3 Pembagian Data Latih dan Data Uji	31
Tabel 4.4 Hasil klasifikasi Naïve bayes	32
Tabel 4.5 Confusion Matrix test_size 0.20	32
Tabel 4.6 Confusion Matrix test_size 0.30	34



DAFTAR GAMBAR

Gambar 3.1 Alur Penelitian	17
Gambar 4.1 code web scrapping	21
Gambar 4.2 Code dari pelabelan.....	23
Gambar 4.3 Dataset Labelling	24
Gambar 4. 4 Code Variabel yang dipakai saja.....	24
Gambar 4.5 Dataset setelah dilakukan Data Cleaning	25
Gambar 4.6 Code dari Case Folding.....	25
Gambar 4.7 Hasil Case Folding	26
Gambar 4.8 Code dari stopword removal	26
Gambar 4.9 Stopword Removal.....	27
Gambar 4.10 Code dari tokenizing	27
Gambar 4.11 Pemisahan teks menjadi potongan	28
Gambar 4.12 Code Stemming	28
Gambar 4.13 Stemming	29
Gambar 4.14 Code untuk Count Vectorizer.....	29
Gambar 4.15 Mengubah matriks fitur frekuensi kata	30
Gambar 4.16 Hasil matriks fitur TF-IDF	30
Gambar 4.17 Splitting Data	31
Gambar 4.18 Visualisasi hasil pengujian	35

INTISARI

Di era digital saat ini, aplikasi mobile semakin banyak digunakan oleh banyak pengguna. Oleh karena itu, penting untuk memahami kepuasan dan ketidakpuasan pengguna saat menggunakan aplikasi tersebut. Zalora merupakan aplikasi e-commerce yang melayani berbagai kebutuhan seperti pakaian, obat-obatan, furniture, mainan dan produk elektronik lainnya. Analisis sentimen ulasan aplikasi e-commerce Zalora di Google Play Store dilakukan dengan menggunakan klasifikasi algoritma Naive Bayes.

Naive Bayes merupakan algoritma yang mempelajari probabilitas suatu objek dengan karakteristik tertentu termasuk dalam kelompok/kelas tertentu. Tujuan dari analisis ini adalah untuk mengetahui hasil analisis dari pendapat pengguna dan melihat kinerja dari algoritma naive bayes. Untuk melihat hasil analisis dari algoritma ini maka dilakukan analisis dengan menggunakan *training* dan *testing* dengan rasio 80% : 20% dan 70% : 30% yang bertujuan menilai data tersebut menghasilkan nilai akurasi pada pengujian yang dilakukan.

Berdasarkan hasil pengujian yang menggunakan data sebanyak 2.500 data, menggunakan algoritma naive bayes. Diperoleh Hasil dengan rasio 80% : 20% mendapatkan akurasi sebesar 82,25% dan rasio 70% : 30% sebesar 81,67%. Dari hasil uji coba tersebut membuktikan bahwa hasil dari kedua rasio ini tidak jauh berbeda tetapi hasil dari rasio 80% : 20% memiliki akurasi yang lebih tinggi daripada rasio 70% : 30%.

Kata kunci: E-Commerce, Zalora, Algoritma Naive Bayes

ABSTRACT

In today's digital era, mobile applications are increasingly being used by many users. Therefore, it is important to understand user satisfaction and dissatisfaction when using the application. Zalora is an e-commerce application that serves various needs such as clothing, medicine, furniture, toys and other electronic products. Sentiment analysis of reviews of the Zalora e-commerce application on the Google Play Store was carried out using the Naive Bayes algorithm classification.

Naive Bayes is an algorithm that studies the probability that an object with certain characteristics belongs to a certain group/class. The purpose of this analysis is to find out the results of the analysis of user opinions and see the performance of the Naive Bayes algorithm. To see the results of the analysis of this algorithm, an analysis was carried out using training and testing with a ratio of 80% : 20% and 70% : 30% which aims to assess the data to produce accuracy values in the tests carried out.

Based on test results using 2,500 data, using the Naive Bayes algorithm. Results obtained with a ratio of 80% : 20% obtained an accuracy of 82.25% and a ratio of 70% : 30% of 81.67%. The test results prove that the results of these two ratios are not much different, but the results of the 80% : 20% ratio have higher accuracy than the 70% : 30% ratio.

Keyword: E-Commerce, Zalora, Naive Bayes Algorithm