



**ANALISIS SENTIMEN OPINI TERHADAP NOVEL PADA WEBSITE  
GOODREADS MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER**

**LAPORAN PENELITIAN**

**RAHMA ZAHRANI  
201410093**

**PROGRAM STUDI SISTEM INFORMASI  
FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS BINA DARMA  
2024**



**ANALISIS SENTIMEN OPINI TERHADAP NOVEL PADA WEBSITE  
GOODREADS MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER**

**RAHMA ZAHRANI  
201410093**

**Laporan Penelitian ini diajukan sebagai syarat memperoleh gelar  
Sarjana Komputer**

**PROGRAM STUDI SISTEM INFORMASI  
FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS BINA DARMA  
2024**

**HALAMAN PENGESAHAN**

**Analisis Sentimen Opini Terhadap Novel Pada Website  
Goodreads Menggunakan Metode Naive Bayes Classifier**

**RAHMA ZAHRANI**

**201410093**

**Telah diterima sebagai salah satu syarat untuk memperoleh gelar  
Sarjana Komputer pada Program Studi Sistem Informasi**

Pembimbing

**Palembang, 25 Juli 2024**


Fakultas Sains Teknologi

Universitas Bina Darma

Dekan,



Universitas Bina Darma  
Fakultas Sains Teknologi



**Nita Rosa Damayanti, M.Kom., Ph.D.**

**Dr. Tata Sutabri, S.Kom., MMSI., MKM**

## HALAMAN PERSETUJUAN

Skripsi Berjudul "Analisis Sentimen Opini Terhadap Novel Pada Website Goodreads Menggunakan Metode Naive Bayes Classifier" Oleh "Rahma Zahrani", telah dipertahankan di depan komisi penguji pada hari Kamis tanggal 25 Juli 2024.

### Komisi Penguji

1. Ketua : Nita Rosa Damayanti, M.Kom., Ph.D.

(.....)

2. Anggota : Evi Yulianingsih, M.M., M.Kom.

(.....)

3. Anggota : Muhamad Ariandi, M.Kom.

(.....)

Mengetahui,  
Program Studi Sistem Informasi  
Fakultas Sains Teknologi  
Universitas Bina Darma  
Ketua,

Universitas Bina  
Darma  
Fakultas Sains Teknologi

(.....)  
Nita Rosa Damayanti, M.Kom., Ph.D.

## SURAT PERNYATAAN

Saya yang bertanda tangan di bawah ini:

Nama : Rahma Zahrani  
NIM : 201410093

Dengan ini menyatakan bahwa:

1. Karya tulis saya (Skripsi) adalah asli dan belum pernah diajukan untuk mendapatkan gelar akademik (Sarjana) di Universitas Bina Darma atau perguruan tinggi lainnya ;
2. Karya tulis ini murni gagasan, rumusan dan penelitian saya dengan arahan dari tim pembimbing ;
3. Di dalam karya tulis ini tidak terdapat karya atau pendapat yang telah ditulis atau dipublikasikan orang lain, kecuali secara tertulis dengan jelas dikutip dengan mencantumkan nama pengarang dan memasukkan ke dalam daftar rujukan ;
4. Saya bersedia tugas skripsi, dicek keasliannya menggunakan plagiarism checker serta diunggah ke internet, sehingga dapat diakses secara daring ;
5. Surat pernyataan ini saya tulis dengan sungguh-sungguh dan apabila terbukti melakukan penyimpangan atau ketidakbenaran dalam pernyataan ini maka saya bersedia menerima sanksi dengan peraturan dan perundang-undangan yang berlaku ;

Demikian surat pernyataan ini saya buat agar dapat dipergunakan sebagaimana mestinya.

Palembang, 25 Juli 2024  
Yang membuat pernyataan,



Rahma Zahrani  
NIM: 201410093

## MOTTO DAN PERSEMBAHAN

### MOTTO

*"Allah tidak membebani seseorang melainkan sesuai dengan kesanggupannya"*

*(QS. Al-Baqarah:286)*

*"It's not always easy, but that's life. Be strong because there are better days ahead"*

*(Mark Lee)*

*"If you never bleed, you're never gonna grow"*

*(Taylor Swift)*

### PERSEMBAHAN

Skripsi ini saya persembahkan kepada:

1. Kedua orang tua saya tercinta, Bapak Wasikin dan Ibu Rumini yang selalu memberikan do'a dan cinta yang tak putus kepada anak-anaknya.
2. Saudara-saudara saya yang selalu setia mendukung dan memotivasi saya.
3. Teman dan sahabat saya.
4. Seluruh dosen Prodi Sistem Informasi.
5. Almamater saya Universitas Bina Darma.

## ABSTRAK

Pertumbuhan teknologi yang pesat memudahkan seseorang dalam menyampaikan berbagai opini secara *online*, termasuk mengenai pengalaman membaca buku. Banyak pembaca ingin mengetahui apakah suatu buku menarik menurut pandangan orang lain sebelum memutuskan buku mana yang akan dibaca. Namun, banyaknya komentar dan opini pada suatu novel mengakibatkan seseorang kesulitan untuk membaca satu per satu opini dan mengambil kesimpulan mengenai novel tersebut. Penelitian ini bertujuan untuk mengetahui kecenderungan publik dan memberikan gambaran mengenai novel "Bumi" berdasarkan opini di *website* Goodreads, serta untuk mengetahui performa Naïve Bayes dalam memberikan klasifikasi sentimen. Metode yang digunakan untuk proses analisis sentimen adalah Naïve Bayes dengan menggabungkan metode Random Oversampling dalam proses *resampling* data latih. Klasifikasi Naïve Bayes dalam memberikan sentimen mendapatkan performa yang cukup baik pada nilai akurasi yakni sebesar 79% dan opini pembaca terhadap novel "Bumi" pada *website* Goodreads cenderung positif dengan persentase 85.26% opini positif, serta berdasarkan visualisasi *wordcloud* pada opini novel "Bumi" dapat disimpulkan novel tersebut bertema fantasi dan mempunyai alur yang lambat bagi sebagian orang.

**Kata Kunci:** Goodreads, Naïve Bayes, Opini, Random Oversampling



## ABSTRACT

The rapid growth of technology has facilitated the sharing of various opinions online, including experiences with reading books. Many readers want to know if a book is interesting based on others' perspectives before deciding which book to read. However, the vast number of comments and opinions on a novel can make it difficult for someone to read each opinion and draw conclusions about the novel. This study aims to understand public tendencies, provide an overview of the novel "Bumi" based on opinions from the Goodreads website, and evaluate the performance of the Naïve Bayes algorithm in sentiment classification. The method used for the sentiment analysis process is Naïve Bayes, combined with the Random Oversampling method for resampling the training data. The Naïve Bayes classification achieved a solid performance in sentiment analysis, with an accuracy rate of 79%. Reader opinions on the novel "Bumi" on Goodreads tend to be positive, with 85.26% positive opinions. Based on the word cloud visualization of the "Bumi" novel opinions, it can be concluded that the novel is fantasy-themed and has a slow pace for some readers.

**Keywords:** Goodreads, Naïve Bayes, Opinion, Random Oversampling



## KATA PENGANTAR

*Bismillahirrahmanirrahim*

Puji dan syukur kepada Tuhan Yang Maha Esa atas limpahan rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan laporan penelitian dengan judul “Analisis Sentimen Opini Terhadap Novel Pada Website Goodreads Menggunakan Metode Naïve Bayes Classifier” dengan baik. Shalawat serta salam tidak lupa selalu tercurahkan kepada Nabi Muhammad SAW, sebagai panutan seluruh umat.

Adapun dalam penyelesaian tugas akhir ini tentunya peneliti mendapatkan banyak saran, motivasi, dorongan hingga bimbingan dari berbagai pihak. Oleh karena itu, dengan segala hormat dan kerendahan hati, ucapan terima kasih secara khusus penulis berikan kepada:

1. Prof. Sunda Ariana, M.Pd., M.M, selaku Rektor Universitas Bina Darma.
2. Dr. Tata Sutabri, S.Kom., MMSI., MKM, selaku Dekan Fakultas Sains Teknologi Universitas Bina Darma.
3. Nita Rosa Damayanti, M.Kom., Ph.D, selaku Ketua Program Studi Sistem Informasi serta dosen pembimbing penelitian yang telah banyak memberikan arahan dan masukan selama proses penelitian.
4. Kurniawan Zaini, M.M., M.Kom, selaku dosen pembimbing akademik penulis.
5. Kedua orang tua penulis yang tersayang, Bapak Wasikin dan Ibu Rumini. Terima kasih banyak telah selalu memberikan semangat, mendoakan, dan selalu memberikan dukungan baik secara moril maupun materiil yang tiada henti-hentinya sehingga penulis dapat menyelesaikan penelitian ini.
6. Kepada kakak-kakak penulis, Adi Raharjo, Linawati, Agustriyanah, S.ST, Kasyfillah, S.IP, Siti Nurdiana, S.AP yang tersayang dan tercinta. Terima kasih telah memberikan dukungan dalam bentuk apapun selama proses penulis menempuh pendidikan ini.

7. Kepada adik penulis, Bagas Ary Pamungkas. Terima kasih telah memberikan dukungan dan do'a yang tiada hentinya selama proses penelitian berlangsung.
8. Kepada rekan-rekan seperjuangan penulis, Umi Kalsum, Alia Qonita, Clarisa Triana, Mayang Puspa, dan Tri Julianti. Terima kasih telah selalu kebersamai penulis selama penelitian, selalu memberikan dukungan satu sama lain, serta menjadi pendengar yang baik bagi penulis selama perkuliahan. Terima kasih atas kebersamaan suka dan duka serta kebaikannya.
9. Terima kasih kepada Taylor Swift, melalui karya-karyanya yang luar biasa selalu menemani penulis selama proses penelitian berlangsung.
10. Terima kasih kepada Nadia Omara dan Nessie Judge, dengan *storytelling*-nya yang selalu menghibur penulis saat penulis dalam keadaan jenuh selama proses penelitian.
11. Kepada Lee Haechan dan seluruh member NCT Dream, selaku idola yang memberikan semangat secara tidak langsung melalui karya-karyanya bagi penulis untuk menyelesaikan penelitian ini.
12. Kepada seluruh pihak yang turut membantu memberikan masukan, motivasi, serta dukungan kepada penulis secara langsung maupun tidak langsung yang tidak dapat penulis sebutkan satu persatu. Semoga Allah membalas ketulusan hati dan kebaikan yang telah diberikan kepada penulis.
13. Terakhir, kepada diri sendiri. Terima kasih telah berjuang sekeras ini dan telah bertahan sampai detik ini. Terima kasih telah memutuskan untuk tidak menyerah dalam menyusun penelitian ini sesulit apapun dengan menyelesaikan pendidikan sebaik dan semaksimal mungkin. *You did your best.*

Penulis menyadari bahwa dalam penyusunan penelitian ini masih terdapat kekurangan. Oleh karena itu, kritik dan saran yang membangun sangat penulis harapkan demi kesempurnaan penelitian ini. Semoga

penelitian ini dapat digunakan sebagaimana mestinya serta bermanfaat bagi semua pihak yang membutuhkan. Aamiin ya rabbal alamin...

Palembang, 25 Juli 2024

Penulis



## DAFTAR ISI

<b>HALAMAN PENGESAHAN .....</b>	<b>ii</b>
<b>HALAMAN PERSETUJUAN .....</b>	<b>iii</b>
<b>SURAT PERNYATAAN .....</b>	<b>iv</b>
<b>MOTTO DAN PERSEMBAHAN .....</b>	<b>v</b>
<b>ABSTRAK .....</b>	<b>vi</b>
<b>ABSTRACT .....</b>	<b>vii</b>
<b>KATA PENGANTAR .....</b>	<b>viii</b>
<b>DAFTAR ISI .....</b>	<b>xi</b>
<b>DAFTAR GAMBAR .....</b>	<b>xiii</b>
<b>DAFTAR TABEL .....</b>	<b>xv</b>
<b>BAB I PENDAHULUAN .....</b>	<b>1</b>
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	4
1.3 Batasan Masalah .....	4
1.4 Tujuan Penelitian .....	4
1.5 Manfaat Penelitian .....	4
<b>BAB II TINJAUAN PUSTAKA .....</b>	<b>6</b>
2.1 Novel .....	6
2.1.1 Pengertian Novel .....	6
2.1.2 Ciri-Ciri Novel .....	6
2.1.3 Jenis-Jenis Novel .....	7
2.2 Situs Goodreads .....	7
2.3 Natural Language Processing (NLP) .....	8
2.4 Machine Learning .....	9
2.5 Analisis Sentimen .....	11
2.6 Text Mining .....	13
2.7 Naïve Bayes .....	15
2.8 Random Oversampling .....	18
2.9 Term Frequency-Inverse Document Frequency (TF-IDF) .....	19

2.10 Confusion Matrix .....	20
2.11 Penelitian Terdahulu .....	22
<b>BAB III METODOLOGI PENELITIAN .....</b>	<b>31</b>
1.1 Pendekatan Penelitian .....	31
1.2 Variabel Penelitian .....	31
1.3 Alur penelitian .....	31
1.4 Pengumpulan Data .....	32
1.5 Pre-processing Data .....	33
1.6 Transformasi Teks .....	35
1.7 Analisis Teks .....	36
1.8 Evaluasi dan Hasil .....	37
<b>BAB IV HASIL DAN PEMBAHASAN .....</b>	<b>38</b>
4.1 Pengumpulan Data .....	38
4.2 Pre-Processing Data .....	44
4.3 Transformasi Teks TF-IDF .....	56
4.4 Analisis Teks Naïve Bayes .....	58
4.5 Evaluasi dan Hasil .....	61
4.6 Visualisasi Data .....	65
<b>BAB V KESIMPULAN DAN SARAN .....</b>	<b>67</b>
1.1 Kesimpulan .....	67
1.2 Saran .....	67
<b>DAFTAR PUSTAKA .....</b>	<b>68</b>
<b>LAMPIRAN .....</b>	<b>72</b>

## DAFTAR GAMBAR

<b>Gambar 2.1</b> Halaman Utama Goodreads.....	8
<b>Gambar 2.2</b> Kategori Machine Learning .....	10
<b>Gambar 2.3</b> Proses Text Mining .....	14
<b>Gambar 3.1</b> Alur Penelitian.....	32
<b>Gambar 4.1</b> Halaman Novel “Bumi” di Website Goodreads .....	39
<b>Gambar 4.2</b> Pemasangan Extension Instant Data Scrapper .....	40
<b>Gambar 4.3</b> Halaman yang Terdeteksi Untuk Pengambilan Data.....	41
<b>Gambar 4.4</b> Data yang Telah Terambil Instant Data Scrapper.....	41
<b>Gambar 4.5</b> Dataset Opini Novel “Bumi” .....	42
<b>Gambar 4.6</b> Grafik Label Sentimen Pada Opini.....	43
<b>Gambar 4.7</b> Kode Program Proses Mengimpor Library dan Casefolding.....	44
<b>Gambar 4.8</b> Kode Program Proses Cleaning Data .....	45
<b>Gambar 4.9</b> Kode Program Menyimpan Data Setelah Proses Casefolding.....	45
<b>Gambar 4.10</b> Kode Program Mengimpor Library Untuk Tokenize .....	46
<b>Gambar 4.11</b> Kode Program Proses Tokenize.....	47
<b>Gambar 4.12</b> Kata-Kata yang Melalui Proses Normalisasi .....	48
<b>Gambar 4.13</b> Kode Program Proses Normalisasi .....	48
<b>Gambar 4.14</b> Kode Program Proses Menyimpan Hasil Normalisasi .....	49
<b>Gambar 4.15</b> Kode Program Proses Mengimpor Library Stopword.....	50
<b>Gambar 4.16</b> Kode Program Proses Stopword .....	51
<b>Gambar 4.17</b> Kode Program Menyimpan Hasil Proses Stopword.....	51
<b>Gambar 4.18</b> Kode Program Mengimpor Library Sastrawi dan Swifter ...	52
<b>Gambar 4.19</b> Kode Program Mendefinisikan Library Stemmer .....	52
<b>Gambar 4.20</b> Kode Program Proses Stemming.....	54
<b>Gambar 4.21</b> Kode Program Menyimpan Hasil Proses Stemming .....	55
<b>Gambar 4.22</b> Grafik Sentimen Setelah Preprocessing.....	56
<b>Gambar 4.23</b> Kode Program Mengimpor TfidfVectorizer .....	56
<b>Gambar 4.24</b> Kode Program Proses TF-IDF .....	57

<b>Gambar 4.25</b> Kode Program Pembagian Antara Data Latih dan Data Uji.....	59
<b>Gambar 4.26</b> Kode Program Proses Random Oversampling.....	60
<b>Gambar 4.27</b> Visualisasi Confusion Matrix.....	61
<b>Gambar 4.28</b> Hasil Perhitungan Evaluasi Setiap Kelas .....	64
<b>Gambar 4.29</b> Visualisasi Wordcloud Sentimen Positif.....	65
<b>Gambar 4.30</b> Visualisasi Wordcloud Sentimen Negatif.....	66





## DAFTAR TABEL

<b>Tabel 2.1</b> Contoh Data Untuk Perhitungan Manual Klasifikasi .....	16
<b>Tabel 2.2</b> Confusion Matrix .....	21
<b>Tabel 2.3</b> Penelitian Terdahulu .....	24
<b>Tabel 4.1</b> Contoh Data Opini .....	42
<b>Tabel 4.2</b> Hasil Proses Cleaning dan Case Folding.....	45
<b>Tabel 4.3</b> Hasil Proses Tokenizing.....	47
<b>Tabel 4.4</b> Hasil Proses Normalisasi .....	49
<b>Tabel 4.5</b> Hasil Proses Filtering .....	51
<b>Tabel 4.6</b> Hasil Proses Stemming.....	55
<b>Tabel 4.7</b> Hasil Proses TF-IDF .....	58
<b>Tabel 4.8</b> Hasil Resampling.....	60
<b>Tabel 4.9</b> Hasil Prediksi Naïve Bayes.....	60