

**ANALISA SENTIMEN DATA ULASAN PENGUNJUNG OBJEK
WISATA GUNUNG BROMO PADA SITUS TRIP ADVISOR
MENGUNAKAN ALGORITMA NAIVE BAYES**



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

SKRIPSI

Alvin Lukita

0000044872

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2024**

**ANALISA SENTIMEN DATA ULASAN PENGUNJUNG OBJEK
WISATA GUNUNG BROMO PADA SITUS TRIP ADVISOR
MENGUNAKAN ALGORITMA NAIVE BAYES**



Diajukan sebagai Salah Satu Syarat untuk Memperoleh
Gelar Sarjana Komputer (S.Kom)

Alvin Lukita
00000044872

PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2024

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Alvin Lukita

Nomor Induk Mahasiswa : 00000044872

Program studi : Sistem Informasi

Skripsi dengan judul:

“Analisa Sentimen Data Ulasan Pengunjung Objek Wisata Gunung Bromo Pada Situs Trip Advisor Menggunakan Algoritma Naive Bayes”

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, 22 Mei 2024

UMIN



(Alvin Lukita)

UNIVERSITAS
MULTIMEDIA
NUSANTARA

HALAMAN PENGESAHAN

Skripsi dengan judul

“Analisa Sentimen Data Ulasan Pengunjung Objek Wisata Gunung Bromo Pada Situs Trip Advisor Menggunakan Algoritma Naïve Bayes”

Oleh

Nama : Alvin Lukita
NIM : 00000044872
Program Studi : Sistem Informasi
Fakultas : Teknik dan Informatika

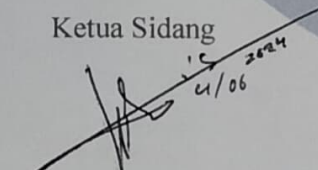
Telah diujikan pada hari Rabu, 22 Mei 2024

Pukul 10.00 s.d 12.00 dan dinyatakan

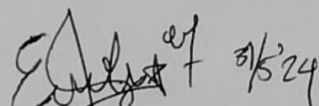
LULUS

Dengan susunan penguji sebagai berikut.

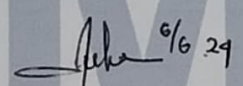
Ketua Sidang


Rudi Sutomo, S.Kom., M.Si., M.Kom.
0222057501

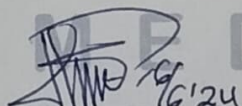
Penguji


Monika Evelin Johan, S.Kom., M.M.S.I.
0327059501/071281

Pembimbing


Melissa Indah Fianty, S.Kom., M.MSI.
0313019201

Ketua Program Studi Sistem Informasi


Ririn Ikana Desanti, S.Kom., M.Kom.
0313058001

iv

Analisa Sentimen Data,...Alvin Lukita, Universitas Multimedia Nusantara

HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH MAHASISWA

Yang bertanda tangan di bawah ini:

Nama : Alvin Lukita
NIM : 00000044872
Program Studi : Sistem Informasi
Jenjang : S1
Judul Karya Ilmiah :

**Analisa Sentimen Data Ulasan Pengunjung Objek Wisata Gunung
Bromo Pada Situs Trip Advisor Menggunakan Algoritma Naive Bayes**

Menyatakan dengan sesungguhnya bahwa saya bersedia (**pilih salah satu**):

- Saya bersedia memberikan izin sepenuhnya kepada Universitas Multimedia Nusantara untuk mempublikasikan hasil karya ilmiah saya ke dalam repositori Knowledge Center sehingga dapat diakses oleh Sivitas Akademika UMN/Publik. Saya menyatakan bahwa karya ilmiah yang saya buat tidak mengandung data yang bersifat konfidensial. Saya tidak akan mencabut kembali izin yang telah saya berikan dengan alasan apapun.
- Saya tidak bersedia mempublikasikan hasil karya ilmiah ini ke dalam repositori Knowledge Center, dikarenakan: Dalam proses pengajuan penerbitan ke dalam jurnal/konferensi nasional/internasional (dibuktikan dengan *letter of acceptance*).

Tangerang, 5 Juni 2024



U M N
UNIVERSITAS
MULTIMEDIA
NUSANTARA
(Alvin Lukita)

* Jika tidak bisa membuktikan LoA jurnal/HKI, saya bersedia mengizinkan penuh karya ilmiah saya untuk dipublikasikan ke KC UMN dan menjadi hak institusi UMN.

KATA PENGANTAR

Puji Syukur atas selesainya penulisan Tugas Akhir ini dengan judul “Analisa Sentimen Data Ulasan Pengunjung Objek Wisata Gunung Bromo Pada Situs Trip Advisor Menggunakan Algoritma Naive Bayes” dilakukan untuk memenuhi salah satu syarat untuk mencapai gelar Sarjana Komputer (S.Kom) Jurusan Sistem Informasi pada Fakultas Teknik dan Informatika Universitas Multimedia Nusantara. Saya menyadari bahwa tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan tugas akhir ini, sangatlah sulit bagi saya untuk menyelesaikan tugas akhir. Oleh karena itu, saya mengucapkan terima kasih kepada:

1. Bapak Dr. Ninok Leksono, selaku Rektor Universitas Multimedia Nusantara.
2. Bapak Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Dekan Fakultas Teknik dan Informatika Universitas Multimedia Nusantara.
3. Ibu Ririn Ikana Desanti, S.Kom., M.Kom., selaku Ketua Program Studi Sistem Informasi Universitas Multimedia Nusantara.
4. Ibu Melissa Indah Fianty, S. Kom., MMSI., sebagai Pembimbing yang telah memberikan bimbingan, arahan, dan motivasi atas terselesainya tugas akhir.
5. Keluarga saya yang telah memberikan bantuan dukungan material dan moral, sehingga penulis dapat menyelesaikan tugas akhir.

Penulis menyadari bahwa penyusunan tugas akhir masih memiliki kekurangan dan jauh dari kata sempurna. Oleh karena itu, penulis menerima kritik, saran yang membangun dari berbagai pihak. Semoga karya ilmiah ini dapat bermanfaat baik sebagai sumber informasi atau sumber inspirasi bagi para pembaca.

Tangerang, 15 Mei 2024

UNIVERSITAS
MULTIMEDIA
NUSANTARA


Alvin Lukita

**ANALISA SENTIMEN DATA ULASAN PENGUNJUNG
OBJEK WISATA GUNUNG BROMO PADA SITUS TRIP
ADVISOR MENGGUNAKAN ALGORITMA NAIVE BAYES**

(Alvin Lukita)

ABSTRAK

Dalam era globalisasi yang terus berkembang, teknologi informasi menjadi pendorong utama transformasi dalam berbagai sektor kehidupan. Teknologi informasi berperan penting dalam kemajuan sektor pariwisata dengan menawarkan kemudahan akses informasi. Objek wisata Gunung Bromo adalah salah satu destinasi populer di Indonesia yang menarik banyak pengunjung setiap tahunnya. Banyaknya komentar dari pengunjung Objek Wisata Gunung Bromo pada situs Trip Advisor menyebabkan tidak ada penyaringan informasi atau ulasan sehingga pihak pengelola taman nasional gunung bromo tidak mengetahui banyak terdapat ulasan positif atau negatif. Ulasan tersebut dapat dianalisa dengan menggunakan analisis sentimen. Pendekatan analisis sentimen terhadap sektor pariwisata dapat digunakan dalam evaluasi layanan destinasi wisata. Metode yang digunakan dalam penelitian adalah Metode Naive Bayes. Penerapan metode klasifikasi *Naive Bayes* dalam pengklasifikasian sentimen positif, negatif. Pada data ulasan pengunjung objek wisata Gunung Bromo, menghasilkan nilai akurasi sebesar 0.85%. Nilai akurasi tersebut dapat digunakan untuk mengklasifikasikan data ulasan pengunjung Objek Wisata Gunung Bromo. Hasil ulasan masyarakat mengenai objek wisata Gunung Bromo adalah terdapat 1001 ulasan, sebanyak 847 ulasan diklasifikasikan sebagai sentimen positif, dan 154 ulasan diklasifikasikan sebagai sentiment negatif. Hal tersebut menunjukkan bahwa ulasan objek wisata Gunung Bromo banyak terdapat ulasan yang bersifat positif. Hasil K-Fold Cross Validation menggunakan model Naive Bayes menghasilkan tingkat akurasi sebesar 0.84%.

Kata kunci: Analisis sentimen, Gunung Bromo, Klasifikasi, Naive Bayes

**SENTIMENT ANALYSIS OF VISITOR REVIEWS DATA AT THE
MOUNT BROMO TOURISM OBJECT ON THE TRIP ADVISOR
SITE USING THE NAIVE BAYES ALGORITHM**

(Alvin Lukita)

ABSTRACT (English)

In the era of ever-growing globalization, information technology is the main driver of transformation in various sectors of life. Information technology plays an important role in the progress of the tourism sector by offering easy access to information. The Mount Bromo tourist attraction is one of the popular destinations in Indonesia which attracts many visitors every year. The large number of comments from visitors to the Mount Bromo tourist attraction on the Trip Advisor site means that there is no filtering of information or reviews so that the Mount Bromo National Park management does not know whether there are many positive or negative reviews. These reviews can be analyzed using sentiment analysis. The sentiment analysis approach to the tourism sector can be used in evaluating tourist destination services. The method used in the research is the Naive Bayes method. Application of the Naïve Bayes classification method in classifying positive and negative sentiments. In the review data from visitors to the Mount Bromo tourist attraction, it produces an accuracy value of 0.85%. This accuracy value can be used to classify visitor review data at the Mount Bromo tourist attraction. The results of public reviews regarding the Mount Bromo tourist attraction are that there are 1001 reviews, 847 reviews are classified as positive sentiment, and 154 reviews are classified as negative sentiment. This shows that there are many reviews of the Mount Bromo tourist attraction that are positive. The results of classification evaluation using the Naive Bayes model show with an accuracy level of 0.84%.

Keywords: *Classification, Mount Bromo, Naive Bayes, Sentiment Analysis*

DAFTAR ISI

HALAMAN PERNYATAAN TIDAK PLAGIAT	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH MAHASISWA	iv
KATA PENGANTAR	v
ABSTRAK	vi
<i>ABSTRACT (English)</i>	vii
DAFTAR ISI	viii
DAFTAR TABEL	x
DAFTAR GAMBAR	xi
DAFTAR RUMUS	xii
DAFTAR LAMPIRAN	xiii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	5
1.3 Batasan Masalah	5
1.4 Tujuan dan Manfaat Penelitian	6
1.4.1 Tujuan Penelitian	6
1.4.2 Manfaat Penelitian	6
1.5 Sistematika Penulisan	6
BAB II LANDASAN TEORI	8
2.1 Penelitian Terdahulu	8
2.2 Teori Tentang Topik	16
2.2.1 Gunung Bromo	16
2.2.2 Analisis Sentimen	17
2.2.3 Word Cloud	18
2.2.4 TF-IDF	19
2.3 Teori Framework dan Algoritma yang digunakan	19
2.3.1 Data Mining	19
2.3.2 Text Mining	20
2.3.3 CRISP-DM	21

2.3.4 Naive Bayes.....	25
2.4 Tools yang digunakan	26
2.4.1 Microsoft Excel	26
2.4.2 Trip Advisor	27
2.4.3 Jupyter Notebook.....	27
2.4.4 Python	27
2.4.5 Octoparse.....	28
BAB III METODOLOGI PENELITIAN	29
3.1 Gambaran Umum Objek Penelitian.....	29
3.2 Metode Penelitian	29
3.3. Alur Penelitian	30
3.4. Teknik Pengumpulan Data.....	34
3.5. Teknik Pengambilan Sampel.....	35
3.6. Teknik Analisis Data	35
BAB IV ANALISIS DAN HASIL PENELITIAN	36
4.1. Identifikasi Masalah.....	36
4.2. Business Understanding.....	36
4.3. Data Understanding	36
4.4. Data Preparation (Pre-Process Data)	42
4.5. TF-IDF.....	49
4.6. Modeling.....	51
4.7. Evaluation	55
4.8. Deployment	58
4.9. Visualization.....	60
4.10. Hasil dan Pembahasan	63
BAB V SIMPULAN DAN SARAN	64
5.1 Kesimpulan	64
5.2 Saran.....	65
DAFTAR PUSTAKA	66
LAMPIRAN.....	72

DAFTAR TABEL

Tabel 1. 1 Tabel Permasalahan dan Dampak	3
Table 2. 1 Tabel Penelitian Terdahulu	8
Tabel 3. 3 Perbandingan Rapid Miner dan Jupyter Notebook.....	35
Tabel 4. 1 Hasil Pembobotan TF-IDF.....	50
Tabel 4. 2 Hasil Confusion Matrix.....	57



UMMN

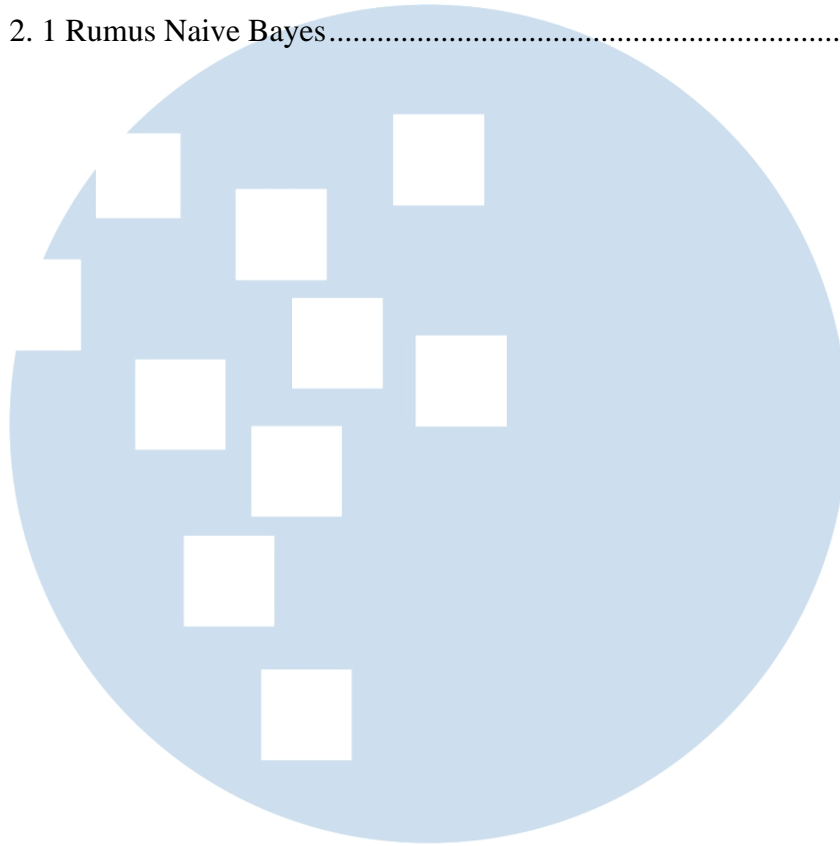
UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR GAMBAR

Gambar 1. 1 Jumlah Wisatawan Gunung Bromo.....	3
Gambar 2. 1 Framework CRISP-DM.....	22
Gambar 3. 1 Alur Penelitian.....	30
Gambar 3. 2 Tahap Data Preparation.....	32
Gambar 4. 1 Data Ulasan Web Scraping.....	37
Gambar 4. 2 Hasil Ulasan Web Scraping.....	37
Gambar 4. 3 Rating Pengunjung.....	38
Gambar 4. 4 Proses Data Labeling.....	39
Gambar 4. 5 Proses Pelabelan Ulasan.....	40
Gambar 4. 6 Hasil Pelabelan.....	40
Gambar 4. 7 Proses plot untuk analisis data.....	41
Gambar 4. 8 Hasil Analisis Data.....	41
Gambar 4. 9 Import Module.....	42
Gambar 4. 10 Import File CSV.....	43
Gambar 4. 11 Proses Case Folding.....	44
Gambar 4. 12 Proses Cleaning.....	45
Gambar 4. 13 Proses review len dan punct.....	45
Gambar 4. 14 Tokenization.....	46
Gambar 4. 15 Filtering Stopword.....	47
Gambar 4. 16 Proses Stemming.....	48
Gambar 4. 17 Proses CountVectorizer.....	49
Gambar 4. 18 Proses Pembobotan TF-IDF.....	50
Gambar 4. 19 Fitur Train Test Split.....	51
Gambar 4. 20 Fitur Ekstrasi.....	52
Gambar 4. 21 Proses K-fold Cross validation.....	54
Gambar 4. 22 Hasil Optimasi GridSearchCV.....	55
Gambar 4. 23 Proses Confusion Matrix.....	56
Gambar 4. 24 Confusion Matrix.....	57
Gambar 4. 25 Uji Coba Prediksi Sentimen.....	58
Gambar 4. 26 Hasil Deployment Website Positif.....	59
Gambar 4. 27 Hasil Deployment Website Negatif.....	59
Gambar 4. 28 Plot Term.....	60
Gambar 4. 29 Proses Visualisasi WordCloud Positif.....	61
Gambar 4. 30 Hasil Visualisasi WordCloud Positif.....	61
Gambar 4. 31 Proses Visualisasi WordCloud Negatif.....	62
Gambar 4. 32 Hasil Visualisasi WordCloud Negatif.....	62

DAFTAR RUMUS

Rumus 2. 1 Rumus Naive Bayes.....	25
-----------------------------------	----

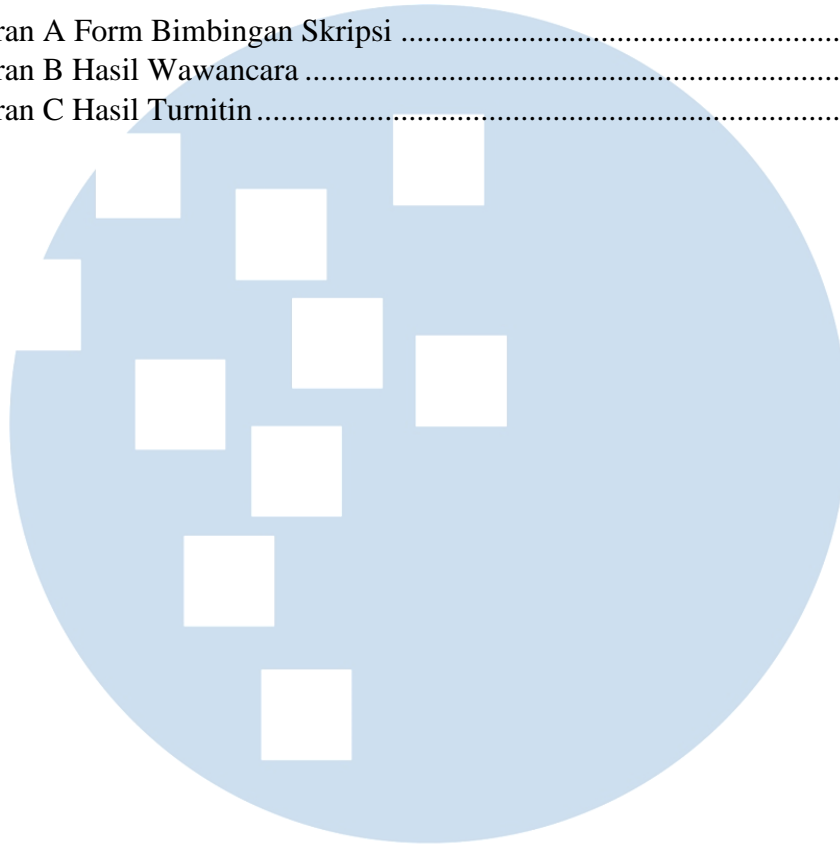


UMMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR LAMPIRAN

Lampiran A Form Bimbingan Skripsi	72
Lampiran B Hasil Wawancara	73
Lampiran C Hasil Turnitin	76



UMMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA