

**ANALISIS SENTIMEN ULASAN TERHADAP TAMAN
MINI INDONESIA INDAH PADA *GOOGLE REVIEW*
MENGUNAKAN METODE *NAÏVE BAYES* DAN
*SUPPORT VECTOR MACHINE***

SKRIPSI

Oleh:

Nida Rachmatin

202010225163



PROGRAM STUDI INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS BHAYANGKARA JAKARTA RAYA

2024

LEMBAR PERSETUJUAN PEMBIMBING

Judul Proposal Tugas Akhir : Analisis Sentimen Ulasan Terhadap Taman Mini
Indonesia Indah Pada *Google Review* Menggunakan
Metode *Naïve Bayes* Dan *Support Vector Machine*

Nama Mahasiswa : Nida Rachmatin

Nomor Pokok Mahasiswa : 202010225163

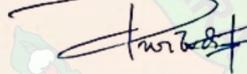
Program Studi/Fakultas : Informatika/Illmu Komputer

Jakarta, 5 Juli 2024

MENYETUJUI,

Pembimbing I

Pembimbing II



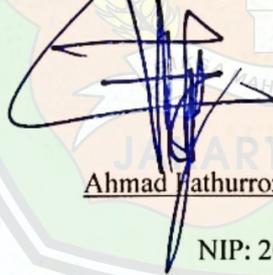
Dr. Herlawati, S.Si., M.M., M.Kom.

Dwi Budi Srisulistiowati., S.Kom., M.M.

NIDN: 0311097302

NIDN: 0323057701

Ketua Program Studi/Sesprodi



Ahmad Athurrozi, S.E., M.M.S.I.

NIP: 2012486

Program Studi Informatika

Fakultas Ilmu Komputer

Universitas Bhayangkara Jakarta Raya

2024

i

LEMBAR PENGESAHAN

Judul Tugas Akhir : Analisis Sentimen Ulasan Terhadap Taman
Mini Indonesia Indah Pada *Google Review*
Menggunakan Metode *Naïve Bayes* Dan
Support Vector Machine

Nama Mahasiswa : Nida Rachmatin

Nomor Pokok Mahasiswa : 202010225163

Program Studi/Fakultas : Informatika/Ilmu Komputer

Tanggal Lulus Ujian Tugas Akhir : 27 Juni 2024

Jakarta, 05 Juli 2024

MENGESAHKAN,

Ketua Tim Penguji : Tri Dharma Putra, S.T., M.Sc.

NIDN : 0302117101

Penguji I : Siti Setiawati, S.Pd., M.Pd.

NIDN : 0313107904

Penguji II : Dr. Herlawati, S.Si., M.M., M.Kom.

NIDN : 0311097302

MENGETAHUI,

Ketua

Dekan

Program Studi Informatika

Fakultas Ilmu Komputer

Ahmad Fathurrozi, S.E., M.M.S.I.
NIP. 2012486

Dr. Dra. Tyastuti Sri Lestari, M.M.
NIP. 1408206



UNIVERSITAS BHAYANGKARA JAKARTA RAYA
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA

LEMBAR PERNYATAAN BUKAN PLAGIASI

Yang bertanda tangan dibawah ini :

Nama : Nida Rachmatin

NPM : 202010225163

Program Studi : Informatika

Fakultas : Ilmu Komputer

Judul Tugas Akhir : Analisis Sentimen Ulasan Terhadap Taman Mini
Indonesia Indah Pada *Google Review* Menggunakan Metode *Naive Bayes* Dan
Support Vector Machine

Dengan ini menyatakan bahwa hasil penulisan skripsi yang telah saya buat ini merupakan **hasil karya saya sendiri dan benar keasliannya**. Apabila dikemudian hari penulisan skripsi ini merupakan plagiat atau penjiplakan terhadap karya orang lain, maka saya bersedia mempertanggungjawabkan sekaligus bersedia menerima sanksi berdasarkan tata tertib di Universitas Bhayangkara Jakarta Raya.

Demikian pernyataan ini saya buat dalam keadaan sadar dan tidak dipaksakan dari pihak manapun.

Jakarta, 05 Juli 2024

Penulis



Nida Rachmatin

ABSTRAK

Nida Rachmatin. 202010225163. Analisis Sentimen Ulasan Terhadap Taman Mini Indonesia Indah Pada *Google Review* Menggunakan Metode *Naïve Bayes* dan *Support Vector Machine*. Informatika. Ilmu Komputer. Universitas Bhayangkara Jakarta Raya. 2024.

Destinasi wisata di Indonesia menawarkan keindahan alam yang menarik minat wisatawan untuk berkunjung. Taman Mini Indonesia Indah (TMII) merupakan salah satu destinasi edukasi dan budaya. Pada tahun 2023, TMII mengalami revitalisasi, Namun, terdapat beberapa masalah terkait fasilitas dan pelayanan yang belum diperbaiki sepenuhnya. Penelitian ini dilakukan untuk mengevaluasi ulasan terhadap TMII pada *Google Review* sebagai bahan evaluasi perbaikan fasilitas pada transportasi dan pelayanan pembayaran tiket masuk dengan pembayaran elektronik dalam meningkatkan kepuasan pengunjung, serta melakukan analisis sentimen terhadap data ulasan TMII yang sudah ada pada ulasan masyarakat pada platform *Google Review*. Selain itu, diolah menggunakan algoritma *Naïve Bayes* dan *Support Vector Machine* (SVM). Hasil analisis menunjukkan bahwa dari total 1.616 data ulasan yang dianalisis, 1.263 ulasan memiliki sentimen negatif dan 353 ulasan memiliki sentimen positif dengan menggunakan metode CRISP-DM. Pengujian model pada dataset yang dibagi dengan rasio 80:20 menunjukkan bahwa model SVM lebih unggul dengan memiliki *accuracy* 85%, *precision* 79%, *recall* 96%, dan *f1-score* 87% sedangkan *Naïve Bayes* memiliki *accuracy* 82%, *precision* 78%, *recall* 90%, dan *f1-score* 83%. Hasil analisis sentimen ini dapat diimplementasikan sebagai sistem untuk memahami dan menganalisis ulasan masyarakat secara efisien, serta sebagai evaluasi untuk meningkatkan fasilitas dan pelayanan di TMII.

Kata Kunci: Analisis Sentimen, *Naïve Bayes*, *Support Vector Machine*, Taman Mini Indonesia Indah, *Google Review*

ABSTRACT

Nida Rachmatin. 202010225163. *Sentiment Analysis of Reviews of Taman Mini Indonesia Indah on Google Reviews Using the Naïve Bayes Method and Support Vector Machine. Informatics. Computer Science. Bhayangkara University, Greater Jakarta. 2024.*

Tourism destinations in Indonesia offer natural beauty that attracts tourists to visit. Taman Mini Indonesia Indah (TMII) is one of the educational and cultural destinations. In 2023, TMII underwent revitalization, however, there are several problems related to facilities and services that have not been fully improved. This research was conducted to evaluate reviews of TMII on Google Review as material for evaluating facility improvements in transportation and entrance ticket payment services with electronic payments in increasing visitor satisfaction, as well as conducting sentiment analysis on TMII review data that already exists in public reviews on the Google Review platform. In addition, it is processed using the Naïve Bayes and Support Vector Machine (SVM) algorithms. The analysis results show that from a total of 1,616 review data analyzed, 1,263 reviews have negative sentiments and 353 reviews have positive sentiments using the CRISP-DM method. Model testing on a dataset divided by a ratio of 80:20 shows that the SVM model is superior by having an accuracy of 85%, precision 79%, recall 96%, and f1-score 87% while Naïve Bayes has an accuracy of 82%, precision 78%, recall 90%, and f1-score 83%. The results of this sentiment analysis can be implemented as a system to understand and analyze community reviews efficiently, as well as an evaluation to improve facilities and services at TMII.

Keywords: *Sentiment Analysis, Naïve Bayes, Support Vector Machine, Taman Mini Indonesia Indah, Google Review*

**LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIK**

Sebagai sivitas akademik Universitas Bhayangkara Jakarta Raya, saya yang bertanda tangan di bawah ini :

Nama : Nida Rachmatin
NPM : 202010225163
Program Studi : Informatika
Fakultas : Ilmu Komputer
Jenis Karya : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Bhayangkara Jakarta Raya **Hak Bebas Royalti Non-Eksklusif (*Non-Exclusive Royalty-Free Right*)**, atas karya ilmiah saya yang berjudul :

Analisis Sentimen Ulasan Terhadap Taman Mini Indonesia Indah Pada *Google Review* Menggunakan Metode *Naive Bayes* Dan *Support Vector Machine*

beserta perangkat yang ada (bila diperlukan). Dengan hak bebas royalti non-eksklusif ini, Universitas Bhayangkara Jakarta Raya berhak menyimpan, mengalihmediakan, mengelolanya dalam bentuk pangkalan data (*database*), mendistribusikannya dan mempublikasikannya di Internet atau media lain untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis dan sebagai pemilik hak cipta.

Segala bentuk tuntutan hukum yang timbul atas pelanggaran hak cipta dalam karya ilmiah ini menjadi tanggung jawab saya pribadi

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta
Pada tanggal : 05 Juli 2024
Yang Menyatakan



Nida Rachmatin

KATA PENGANTAR

Puji syukur kami panjatkan ke hadirat Allah SWT, karena atas rahmat dan hidayah-Nya, kami dapat menyelesaikan penyusunan skripsi ini yang berjudul **“Analisis Sentimen Ulasan Terhadap Taman Mini Indonesia Indah Pada *Google Review* Menggunakan Metode *Naïve Bayes* Dan *Support Vector Machine*”**. Sebagai salah satu syarat yang harus dipenuhi dalam menyelesaikan jenjang strata satu program studi Informatika, Fakultas Ilmu Komputer, Universitas Bhayangkara Jakarta Raya.

Dalam penyusunan skripsi ini, kami menyadari sepenuhnya bahwa selesainya skripsi ini banyak sekali mendapatkan bantuan dukungan, semangat, serta bimbingan dari berbagai pihak, oleh sebab itu penulis ingin mengucapkan rasa terima kasih kepada:

1. Bapak Irjen Pol (Purn) Drs. H. Bambang Karsono, S.H., M.M. selaku Rektor Universitas Bhayangkara Jakarta Raya.
2. Ibu Dr. Tyastuti Sri Lestari, S.Si., M.M. selaku Kepala Program Studi Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya.
3. Bapak Ahmad Fahturozzi, SE., MMSI, selaku Ketua Program Studi Informatika. Universitas Bhayangkara Jakarta Raya.
4. Ibu Dr. Herlawati, S.Si., M.M., M.Kom. selaku Dosen Pembimbing I Skripsi yang banyak mengajarkan dan memberikan masukan serta saran dan motivasi dalam mengerjakan tugas akhir ini.

5. Ibu Dwi Budi Srisulistiowati., S.Kom., M.M, Selaku Dosen Pembimbing II Skripsi yang banyak mengajarkan dan memberikan masukan serta saran dan motivasi dalam mengerjakan tugas akhir ini.
6. Kedua orang tua tercinta yang selalu mendoakan dengan Ikhlas, memberi dukungan dan semangat kepada penelitian ini.
7. Seluruh dosen Fakultas Ilmu Komputer yang telah mendidik dan memberikan ilmu selama kuliah delapan semester ini dan seluruh staf yang selalu sabar melayani segala administrasi selama proses penelitian ini.
8. Sang pujaan hati Andriansyah selaku *support system* penulis yang selalu memberi semangat selama penulisan skripsi berlangsung.
9. Serta sahabat Prili, Cesa, Silvi yang selalu mendukung dan membantu di dunia perkuliahan dengan penuh liku-liku bersama.

Kami menyadari bahwa skripsi ini masih jauh dari kesempurnaan. Oleh karena itu, segala saran dan kritik yang membangun dari semua pihak sangat kami harapkan demi perbaikan pada skripsi selanjutnya. Harapan kami semoga skripsi ini bermanfaat bagi pembaca lain pada umumnya.

Jakarta, 05 Maret 2024

Hormat saya



Nida Rachmatin

202010225163

DAFTAR ISI

LEMBAR PERSETUJUAN PEMBIMBING	i
LEMBAR PENGESAHAN	ii
LEMBAR PERNYATAAN BUKAN PLAGIASI	iii
ABSTRAK	iv
<i>ABSTRACT</i>	v
LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI.....	vi
KATA PENGANTAR.....	vii
DAFTAR ISI.....	ix
DAFTAR TABEL	xiii
DAFTAR GAMBAR	xiv
DAFTAR LAMPIRAN	xvii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	5
1.3 Rumusan Masalah	5
1.4 Batasan Masalah.....	6
1.5 Tujuan Penelitian.....	6
1.6 Manfaat Penelitian.....	7
1.7 Sistematika Penulisan.....	7

BAB II TINJAUAN PUSTAKA	9
2.1 Penelitian Terdahulu	9
2.2 Landasan Teori	14
2.2.1 <i>Machine Learning</i>	14
2.2.2 <i>Data Mining</i>	15
2.2.3 <i>Cross-Industry Standard Process for Data Mining (CRISP-DM)</i> ...17	
2.2.4 <i>Text Mining</i>	19
2.2.5 Pembobotan TF-IDF	22
2.2.6 Analisis Sentimen	24
2.2.7 Taman Mini Indonesia Indah	25
2.2.8 <i>Google Review</i>	26
2.2.9 <i>Naïve Bayes</i>	26
2.2.10 <i>Support Vector Machine (SVM)</i>	28
2.2.11 <i>Confusion Matrix</i>	30
2.2.12 <i>Lexicon Based</i>	32
2.2.13 SMOTE	33
2.2.14 <i>Python</i>	34
BAB III METODOLOGI PENELITIAN	36
3.1 Objek Penelitian	36

3.2	Kerangka Penelitian	37
3.2.1	<i>Business Understanding</i> (Pemahaman Bisnis).....	37
3.2.2	<i>Data Understanding</i> (Pemahaman Data)	38
3.2.3	<i>Data Preparation</i> (Persiapan Data)	40
3.2.4	<i>Modelling</i> (Pemodelan).....	41
3.2.5	<i>Evaluation</i> (Evaluasi).....	44
3.2.6	<i>Deployment</i> (Implementasi)	44
BAB IV	HASIL DAN PEMBAHASAN.....	45
4.1	<i>Business Understanding</i> (Pemahaman Bisnis).....	45
4.2	<i>Data Understanding</i> (Pemahaman Data).....	46
4.3	<i>Data Preparation</i> (Persiapan Data)	50
4.4	<i>Modelling</i> (Pemodelan)	66
4.4.1	<i>Split Data</i>	66
4.4.2	Hasil <i>Modelling Naïve Bayes</i>	66
4.4.3	Hasil <i>Modelling Support Vector Machine</i>	72
4.5	<i>Evaluation</i> (Evaluasi)	78
4.5.1	<i>Naïve Bayes</i>	78
4.5.2	<i>Support Vector Machine</i>	81
4.5.3	<i>Wordcloud</i>	85

4.6	<i>Deployment (Implementasi)</i>	88
BAB V PENUTUP		93
5.1	Kesimpulan	93
5.2	Saran	94
DAFTAR PUSTAKA		96
LAMPIRAN		105



DAFTAR TABEL

Tabel 2.1 Penelitian Terkait	9
Tabel 4.1 Hasil <i>Case Folding</i>	51
Tabel 4.2 Hasil <i>Tokenizing</i>	52
Tabel 4.3 Hasil Normalisasi.....	53
Tabel 4.4 Hasil <i>Stopword</i>	54
Tabel 4.5 Hasil <i>stemming</i>	55
Tabel 4.6 Contoh dokumen untuk perhitungan TF-IDF	60
Tabel 4.7 Hasil TF Normalisasi	61
Tabel 4.8 Hasil dari <i>Document Frequency (DF)</i>	62
Tabel 4.9 Hasil dari IDF	63
Tabel 4.10 Hasil dari TF-IDF	64
Tabel 4.11 Data <i>Training</i>	67
Tabel 4.12 Prediksi Data <i>Training Naive Bayes</i>	68
Tabel 4.13 Untuk Uji Data <i>Testing</i>	70
Tabel 4.14 Hasil probabilitas pada Data <i>Testing</i>	72
Tabel 4.15 Contoh perhitungan manual SVM	74
Tabel 4.16 Tabel <i>Confusion Matrix Naive Bayes</i>	79
Tabel 4.17 Hasil perbandingan antara <i>Naive Bayes</i> dan SVM	85
Tabel 4.18 Jumlah kata yang sering muncul pada ulasan positif.....	85
Tabel 4.19 Jumlah kata yang sering muncul pada ulasan negatif.....	87

DAFTAR GAMBAR

Gambar 1.1 Ulasan Pengguna <i>Google Review</i> TMII.....	3
Gambar 2.1 <i>Cross-Industry Standard Process for Data Mining</i> (CRISP – DM) .	17
Gambar 2.2 Tahapan data <i>Preprocessing</i>	21
Gambar 2.3 Logo TMII	25
Gambar 2.4 <i>Support Vector Machine</i>	29
Gambar 2.5 <i>Confusion Matrix</i>	30
Gambar 3.1 Peta Lokasi TMII	36
Gambar 3.2 Kerangka Penelitian.....	37
Gambar 3.3 Hasil <i>Scrapping</i> Data Ulasan TMII	38
Gambar 3.4 Teknik <i>Web Scrapping</i>	39
Gambar 3.5 Proses Algoritma <i>Naïve Bayes</i>	41
Gambar 3.6 Proses Algoritma <i>Support Vector Machine</i> (SVM)	43
Gambar 4.1 <i>Instant Data Scraper</i>	46
Gambar 4.2 Halaman <i>google maps</i>	47
Gambar 4.3 <i>Scrapping</i> Data	48
Gambar 4.4 Pengambilan Data	49
Gambar 4.5 Hasil <i>Scrapping</i> Data Ulasan TMII	49
Gambar 4.6 Kode <i>Case Folding</i>	50
Gambar 4.7 Kode <i>Tokenizing</i>	51
Gambar 4.8 Daftar kata kamus <i>Slang Word</i>	52
Gambar 4.9 Kode Normalisasi	53
Gambar 4.10 Kode <i>Stopword</i>	54

Gambar 4.11 Hasil <i>Steaming</i>	55
Gambar 4.12 Kode <i>Labelling</i>	56
Gambar 4.13 <i>Output Total Labelling</i>	57
Gambar 4.14 Grafik data sentimen.....	57
Gambar 4.15 Hasil <i>Labelling</i>	58
Gambar 4.16 Kode TF – IDF Perkalimat	59
Gambar 4.17 Hasil TF-IDF Perkalimat	59
Gambar 4.18 Hasil TF – IDF	60
Gambar 4.19 contoh hasil dan kode SMOTE	65
Gambar 4.20 Hasil dan Grafik Setelah SMOTE	65
Gambar 4.21 Kode <i>Split Data</i>	66
Gambar 4.22 Kode dan Hasil <i>Modelling Naive Bayes</i>	67
Gambar 4. 23 Hasil dan kode SVM.....	73
Gambar 4.24 Grafik Kernel Linear SVM.....	73
Gambar 4.25 Tabel <i>Confusion Matrix Naive bayes</i>	78
Gambar 4.26 Hasil <i>Confusion Matrix Naive Bayes</i>	79
Gambar 4.27 Hasil <i>Confusion Matrix Support Vector Machine</i>	81
Gambar 4.28 Hasil <i>Confusion Matrix SVM</i>	82
Gambar 4.29 <i>Wordcloud</i> Positif.....	86
Gambar 4.30 <i>Wordcloud</i> Negatif	87
Gambar 4.31 Tampilan <i>website</i> untuk Input kalimat Sentimen.....	89
Gambar 4.32 Input kalimat yang berpotensi memiliki sentimen positif	90
Gambar 4.33 Tampilan <i>website</i> untuk Input kalimat Sentimen.....	91

Gambar 4.34 Input kalimat yang berpotensi memiliki sentimen negatif92



DAFTAR LAMPIRAN

Lampiran 1 : Plagiarisme

Lampiran 2 : Biodata Mahasiswa

Lampiran 3 : Kartu Bimbingan Mahasiswa

Lampiran 4 : Data Mentah Ulasan TMII Pada *Google Review*

