



**ANALISIS SENTIMEN TERHADAP PENILAIAN OBJEK WISATA
DI KOTA DKI JAKARTA DENGAN MENERAPKAN ALGORITMA
NAIVE BAYES CLASSIFIER**

SKRIPSI

JONES ANANTA TARIGAN

1810511027

**PROGRAM STUDI SARJANA INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN VETERAN JAKARTA
JAKARTA
2022**



**ANALISIS SENTIMEN TERHADAP PENILAIAN OBJEK WISATA
DI KOTA DKI JAKARTA DENGAN MENERAPKAN ALGORITMA
NAIVE BAYES CLASSIFIER**

SKRIPSI

**Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh Gelar Sarjana
Komputer**

JONES ANANTA TARIGAN

1810511027

**PROGRAM STUDI SARJANA INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN VETERAN JAKARTA
JAKARTA
2022**

PERNYATAAN ORISINALITAS

Tugas Akhir ini adalah hasil karya sendiri, dan semua sumber yang dikutip maupun yang dirujuk telah saya nyatakan dengan benar.

Nama : Jones Ananta Tarigan

NIM : 1810511027

Tanggal : 16 November 2022

Bilamana di kemudian hari ditemukan ketidaksesuaian dengan pernyataan saya ini, maka saya bersedia dituntut dan diproses sesuai ketentuan yang berlaku.

Jakarta, 16 November 2022

Yang Menyatakan



(Jones Ananta Tarigan)

PERNYATAAN PERSETUJUAN PUBLIKASI

Sebagai civitas akademik Universitas Pembangunan Nasional Veteran Jakarta, saya yang bertandatangan di bawah ini:

Nama : Jones Ananta Tarigan

NIM : 1810511027

Fakultas : Ilmu Komputer

Program Studi : S1 Informatika

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

**ANALISIS SENTIMEN TERHADAP PENILAIAN OBJEK WISATA DI
KOTA DKI JAKARTA DENGAN MENERAPKAN ALGORITMA NAIVE
BAYES CLASSIFIER**

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti ini Universitas Pembangunan Nasional Veteran Jakarta berhak menyimpan, mengalih media/formatkan, mengelola dalam bentuk pangkalan kata (Basis data), merawat dan mempublikasikan Skripsi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta. Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta

Pada Tanggal : 16 November 2022

Yang Menyatakan,



(Jones Ananta Tarigan)

LEMBAR PENGESAHAN

Dengan ini dinyatakan bahwa Skripsi berikut :

Nama : Jones Ananta Tarigan

NIM : 181051107

Program Studi : Informatika

Judul Skripsi : Analisis Sentimen Terhadap Penilaian Objek Wisata Di Kota DKI Jakarta Dengan Menerapkan Algoritma Naive Bayes Classifier

Telah berhasil dipertahankan di hadapan Tim Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Komputer pada Program Studi S1 Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta.



Dr. Didit Widiyanto, S.Kom., M.Si.

Penguji I



Iin Ernawati, S.Kom., M.Si.

Penguji II



Theresiawati, S. Kom., M.TI.

Pembimbing I



Sarika, S.Kom., M.Kom.

Pembimbing II



Dr. Ermatita, M.Kom.

Dekan



Dr. Widya Cholil, M.TI.
Ketua Program Studi

Ditetapkan di : Jakarta
Tanggal : 13 Desember 2022



LEMBAR PERSETUJUAN

Dengan ini dinyatakan bahwa Tugas Akhir berikut :

Nama : Jones Ananta Tarigan

NIM : 1810511027

Program Studi : S-1 Informatika

Judul Tugas Akhir : ANALISIS SENTIMEN TERHADAP PENILAIAN OBJEK
WISATA DI KOTA DKI JAKARTA DENGAN MENERAPKAN ALGORITMA NAIVE
BAYES CLASSIFIER

Sebagai bagian persyaratan yang diperlukan untuk mengikuti ujian Sidang Tugas
Akhir/Skripsi pada Program Studi Informatika Fakultas Ilmu Komputer, Universitas
Pembangunan Nasional Veteran Jakarta.

Menyetujui,

Dosen Pembimbing I



(Theresawati, S. Kom., MTI.)

Dosen Pembimbing 2



(Sarika, S.Kom., M.Kom.)

Menyetujui,

(Dr. Widya Cholil, MIT)

Kepala Program Studi

Ditetapkan di : Jakarta

Tanggal Persetujuan : 18 November 2022

ANALISIS SENTIMEN TERHADAP PENILAIAN OBJEK WISATA DI KOTA DKI JAKARTA DENGAN MENERAPKAN ALGORITMA NAIVE BAYES CLASSIFIER

Jones Ananta Tarigan

ABSTRAK

Dengan pesatnya perkembangan teknologi informasi, membawa dampak positif serta keuntungan bagi para pelaku wisata. Pesatnya penggunaan teknologi informasi juga berbanding lurus dengan perkembangan destinasi wisata tersebut. Perkembangan destinasi wisata juga mengimplementasikan kebijakan yang dibuat oleh pemerintah wilayah tersebut untuk keberlangsungan objek wisata. Kritik dan saran dari turis serta wisatawan sangat dibutuhkan dalam peningkatan objek wisata tersebut sehingga kebijakan yang belum optimal di masing – masing objek wisata dapat segera diperbaiki. Dengan adanya *feedback* dari turis maupun wisatawan, dapat mempermudah Pemerintah Daerah DKI Jakarta dalam mengambil keputusan. Penelitian ini dilakukan bertujuan guna menerapkan metode *Naive Bayes Classifier* pada tempat objek wisata di wilayah DKI Jakarta yang telah diberikan komentar atau masukan dari para pengunjung. Data bersumber dari situs *Tripadvisor* dan diperoleh data sebanyak 818. Penelitian ini, mengevaluasi kinerja penerapan metode *Naive Bayes Classifier* agar diterapkan sebagai algoritma dengan nilai akurasi tinggi. Tingkat akurasi akan menggunakan metode confusion matrix beserta tabel presisi dari masing – masing model kriteria sehingga diperoleh nilai untuk kelas positif yakni 91%, bagi kelas negatif yakni 60%, kelas positif negatif yakni 48%.

Kata Kunci: Objek Wisata, *Tripadvisor*, *Naive Bayes Classifier*

ANALYSIS OF TOURISM OBJECTIVE ASSESSMENT IN DKI JAKARTA BY USING THE NAIVE BAYES CLASSIFIER ALGORITHM

Jones Ananta Tarigan

ABSTRACT

With the rapid development of information technology, it has a positive impact and benefits for tourism actors. The rapid use of information technology is also directly proportional to the development of these tourist destinations. The development of tourist destinations also implements policies made by the regional government for the sustainability of tourism objects. Criticisms and suggestions from tourists and tourists are needed in improving these attractions s policies that are not optimal in each tourist attraction can be immediately improve by the government. With feedback from tourists, it can make it easier for the DKI Jakarta Regional Government to make decisions. This study aims to apply the Naive Bayes Classifier method to tourist attractions in the DKI Jakarta area which have been given comments or input from visitors. Data sourced from the Tripadvisor website and gain 818 is obtained. This study evaluates the performance of the application of the *Naive Bayes Classifier* method to be applied as an algorithm with a high accuracy value. The level of accuracy will use the confusion matrix method along with the precision table of each criteria model and gain value for the positive class is 91%, for the negative class is 60%, the positive or negative class that is 48%.

Keyword: *Tourism Objective, Tripadvisor, Naive Bayes Classifier*

KATA PENGANTAR

Puji dan syukur penulis panjatkan sebesar – besarnya kepada Tuhan Yang Maha Esa, oleh karena berkat dan pengasihannya di dalam pengerjaan Skripsi yang berjudul “Analisis Sentimen Terhadap Penilaian Objek Wisata Di Kota DKI Jakarta Dengan Menerapkan Algoritma Naive Bayes Classifier” dapat terselesaikan. Dalam pengerjaan penulisan ini, penulis juga mengucapkan terimakasih terkhusus kepada:

1. Kedua orangtua yang selalu memberikan dukungan yang terbaik serta seluruh keluarga
2. Ibu Theresiawati, S. Kom., MTI. Dan Ibu Sarika, S.Kom., M.Kom. selaku dosen pembimbing yang sudah memberikan saran dan masukan selama pengerjaan skripsi ini
3. Bapak Dr. Didit Widiyanto, S. Kom., M.Si. selaku dosen pembimbing akademik
4. Seluruh dosen dan sivitas akademik Fakultas Ilmu Komputer yang sudah memberikan ilmu dan pengajaran selama menempuh pendidikan
5. Teman – teman perkuliahan yang telah berjuang dalam setiap proses kuliah dan memberikan saran dalam pengerjaan skripsi
6. Setiap pihak yang membantu dalam proses pengerjaan skripsi ini yang tidak dapat penulis sebutkan namanya.

Jakarta, 16 November 2022

Yang Menyatakan



(Jones Ananta Tarigan)

DAFTAR ISI

PERNYATAAN ORISINALITAS	iii
LEMBAR PENGESAHAN	v
LEMBAR PERSETUJUAN	v
ABSTRAK	vii
ABSTRACT	viii
KATA PENGANTAR	ix
DAFTAR ISI	x
DAFTAR GAMBAR	xii
DAFTAR TABEL	xiii
BAB I PENDAHULUAN	14
1.1 Latar Belakang	14
1.2 Rumusan Masalah	16
1.3 Batasan Masalah	16
1.4 Tujuan	16
1.5 Manfaat	16
1.6 Sistematika Penulisan	17
BAB II LANDASAN TEORI	18
2.1 Pariwisata	18
2.2 Machine Learning	18
2.3 Pengolahan Bahasa Alami	19
2.4 Analisis Sentimen	19
2.5 <i>Text Mining</i>	20
2.7 <i>Naïve Bayes Classification</i>	22
2.8 <i>Performance Evaluation Measure</i>	23
2.9 Penelitian Serupa	25
BAB III METODE PENELITIAN	28
3.1 Alur Pengerjaan Tugas Akhir	28
3.2 Uraian Metodologi	29
3.2.1 Pengambilan Data	29
3.2.2 Langkah-Langkah Preprocessing	31
3.2.3 Ekstraksi Fitur	31
3.2.4 Klasifikasi Naive Bayes	31
3.2.5 Uji Model	32
3.2.6 Evaluasi Model	32
BAB IV HASIL DAN PEMBAHASAN	33
4.1 Implementasi	33
4.1.1 Data	33
4.1.2 <i>Labelling</i>	36
4.1.3 <i>Preprocessing</i>	37
4.1.3.1 <i>Cleaning</i>	37
4.1.3.2 <i>Remove Stopword</i>	39
4.1.3.3 <i>Stemming</i>	40
4.1.3.4 <i>Tokenization</i>	41
4.1.4 Pembobotan Term (TF-IDF)	41
4.1.5 Klasifikasi <i>Naive Bayes</i>	47

4.1.5.1 Data Training	47
4.1.5.2 Data <i>Testing</i>	48
4.1.6 Uji Model	50
4.1.7 Evaluasi Model	51
5.1 Kesimpulan	53
5.2 Saran	54
DAFTAR PUSTAKA	55
RIWAYAT HIDUP	57
LAMPIRAN	58

DAFTAR GAMBAR

Gambar 2. 1 Kerangka Kerja	21
Gambar 2.2 Perbandingan Tingkat Akurasi dan Presisi(a), Akurat(b), Presisi(c), Tidak Akurat dan Tidak Presisi(d).	25
Gambar 3. 1 Alur Pengerjaan Tugas Akhir.....	28
Gambar 3. 2 Halaman Awal Situs <i>Tripadvisor</i>	29
Gambar 4.1 Plug-in Instant Data Scraper	33
Gambar 4.2 Ulasan Pengunjung Objek Wisata Masjid Istiqlal	34
Gambar 4.3 Implementasi <i>Instant Data Scraper</i>	35
Gambar 4.4 Implementasi Tahapan <i>Cleaning</i>	38
Gambar 4.5 Implementasi Tahapan <i>Remove Stopword</i>	39
Gambar 4.6 Implementasi Tahapan <i>Remove Stopword</i>	39
Gambar 4.7 Implementasi Tahapan <i>Remove Stopword</i>	40
Gambar 4.8 Tahapan <i>Stemming</i>	41
Gambar 4.9 Tahapan <i>Tokenization</i>	41
Gambar 4.10 Keluaran dari tahapan <i>preprocessing</i>	41
Gambar 4.11 Kode Word Cloud Positif	42
Gambar 4.12 Kode Word Cloud Negatif	42
Gambar 4.13 Source Code Word Cloud Netral	42
Gambar 4.14 Gambaran Visual Positif	43
Gambar 4.15 Gambaran Visual Negatif	43
Gambar 4.16 Gambaran Visual Netral	44
Gambar 4.17 <i>Confusion Matrix</i> dan Nilai Akurasi	51

DAFTAR TABEL

Tabel 2.1 <i>Confusion Matrix</i>	24
Tabel 3.1 Contoh Data Hasil <i>Labelling</i>	30
Tabel 4.1 Data Ulasan Pengunjung Setelah <i>Scraping</i>	36
Tabel 4.2 Tahapan Pembobotan <i>Term</i>	37
Tabel 4.3 Data <i>Testing</i>	48
Tabel 4.4 Penghitungan Data <i>Testing</i>	48
Tabel 4.5 Penghitungan Peluang Data <i>Training</i>	49
Tabel 4.6 Hasil Tahapan Uji Model	50
Tabel 4.7 Kelas Positif	51
Tabel 4.8 Kelas Negatif	51
Tabel 4.9 Kelas Positif Negatif	52