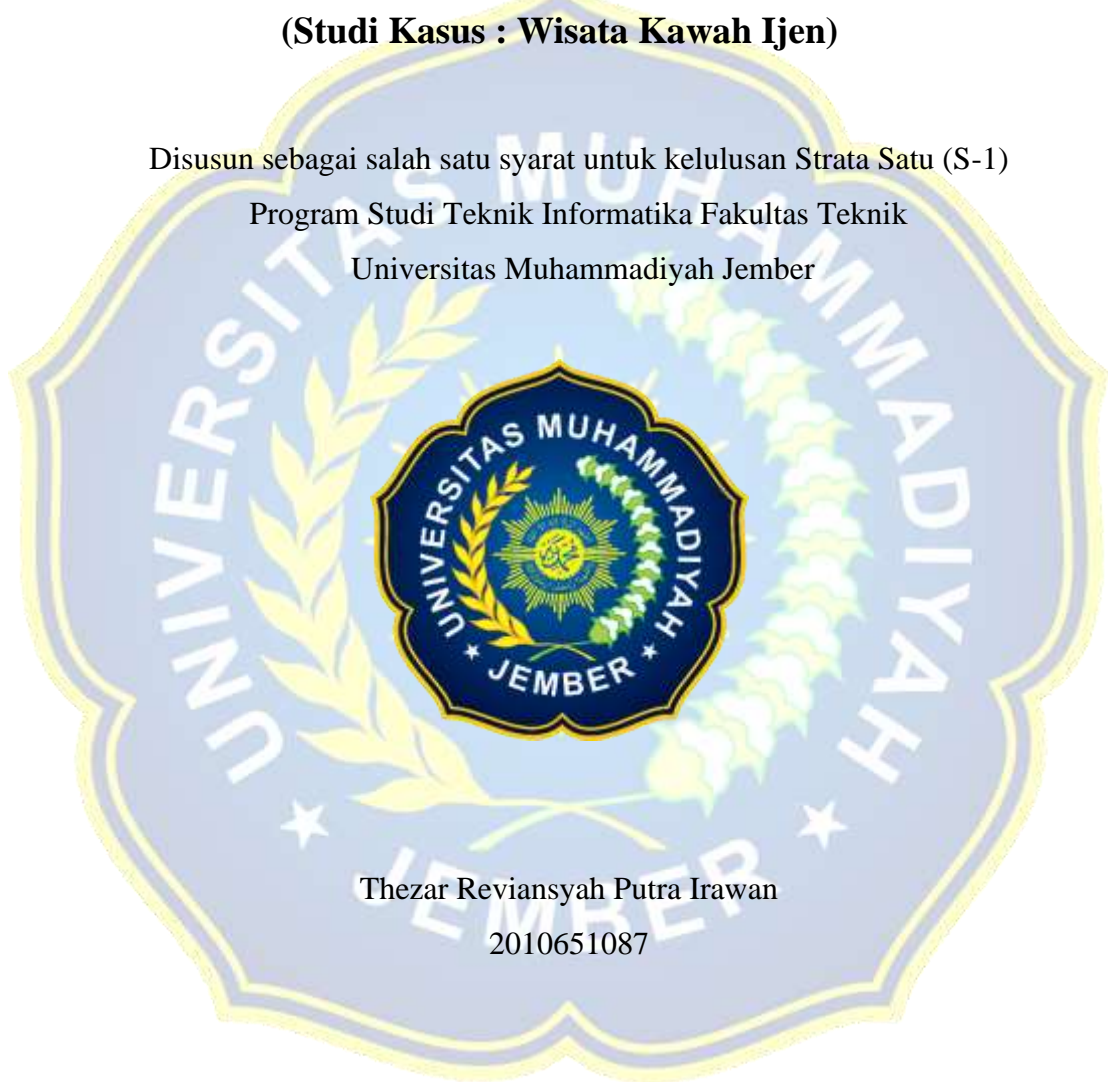


TUGAS AKHIR

ANALISIS SENTIMEN TERHADAP WISATA DI KABUPATEN BANYUWANGI MENGGUNAKAN METODE *NAIVE BAYES* *CLASSIFIER*

(Studi Kasus : Wisata Kawah Ijen)

Disusun sebagai salah satu syarat untuk kelulusan Strata Satu (S-1)
Program Studi Teknik Informatika Fakultas Teknik
Universitas Muhammadiyah Jember



Thezar Reviansyah Putra Irawan
2010651087

**FAKULTAS TEKNIK
PROGRAM STUDI TEKNIK INFORMATIKA
UNIVERSITAS MUHAMMADIYAH JEMBER**

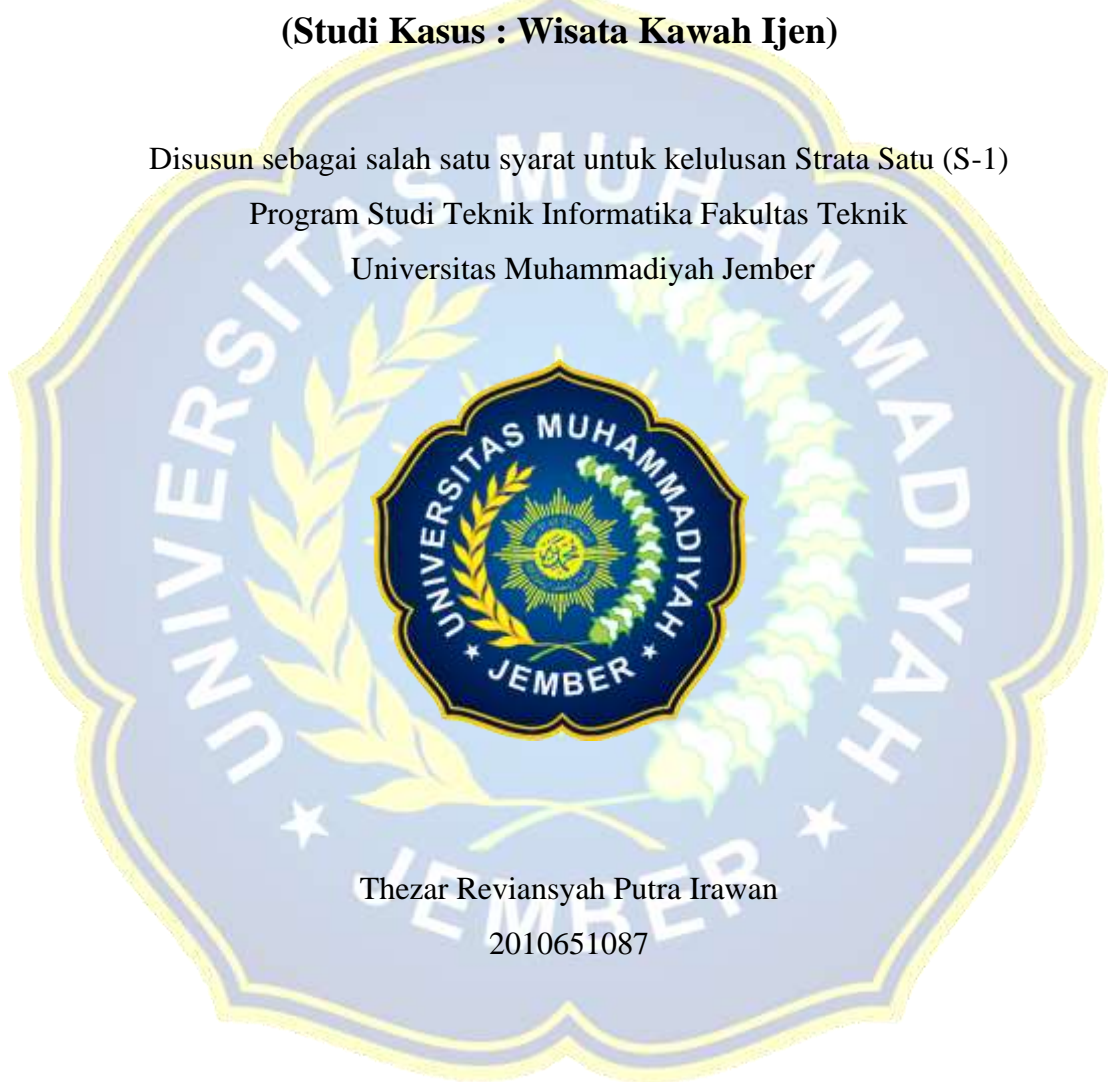
2024

TUGAS AKHIR

ANALISIS SENTIMEN TERHADAP WISATA DI KABUPATEN BANYUWANGI MENGGUNAKAN METODE *NAIVE BAYES* *CLASSIFIER*

(Studi Kasus : Wisata Kawah Ijen)

Disusun sebagai salah satu syarat untuk kelulusan Strata Satu (S-1)
Program Studi Teknik Informatika Fakultas Teknik
Universitas Muhammadiyah Jember



Thezar Reviansyah Putra Irawan
2010651087

**FAKULTAS TEKNIK
PROGRAM STUDI TEKNIK INFORMATIKA
UNIVERSITAS MUHAMMADIYAH JEMBER**

2024

SURAT PERNYATAAN

Yang bertanda tangan di bawah ini :

Nama : Thezar Reviansyah Putra Irawan
Nomor Induk Mahasiswa : 2010651087
Program Studi : Teknik Informatika
Perguruan Tinggi : Universitas Muhammadiyah Jember

Menyatakan dengan sesungguhnya karya ilmiah berupa tugas akhir yang berjudul **“ANALISIS SENTIMEN TERHADAP WISATA DI KABUPATEN BANYUWANGI MENGGUNAKAN METODE *NAIVE BAYES CLASSIFIER* (Studi Kasus : Wisata Kawah Ijen)”** adalah benar merupakan karya sendiri. Hal-hal yang bukan karya saya, dalam skripsi tersebut diberi tanda sitasi dan ditunjukkan dalam daftar pustaka. Apabila di kemudian hari terbukti pernyataan saya tidak benar dan ditemukan pelanggaran atas karya Tugas Akhir ini, saya bersedia menerima sanksi akademik berupa pencabutan Tugas Akhir dan gelar yang saya peroleh dari Tugas Akhir tersebut.

Jember, 29 Juli 2024

Penulis

Thezar Reviansyah P.I

NIM. 2010651087

SURAT PERNYATAAN

Yang bertanda tangan di bawah ini :

Nama : Thezar Reviansyah Putra Irawan
Nomor Induk Mahasiswa : 2010651087
Program Studi : Teknik Informatika
Perguruan Tinggi : Universitas Muhammadiyah Jember

Menyatakan dengan sesungguhnya karya ilmiah berupa tugas akhir yang berjudul **"ANALISIS SENTIMEN TERHADAP WISATA DI KABUPATEN BANYUWANGI MENGGUNAKAN METODE *NAIVE BAYES CLASSIFIER* (Studi Kasus : Wisata Kawah Ijen)"** adalah benar merupakan karya sendiri. Hal-hal yang bukan karya saya, dalam skripsi tersebut diberi tanda sitasi dan ditunjukkan dalam daftar pustaka. Apabila di kemudian hari terbukti pernyataan saya tidak benar dan ditemukan pelanggaran atas karya Tugas Akhir ini, saya bersedia menerima sanksi akademik berupa pencabutan Tugas Akhir dan gelar yang saya peroleh dari Tugas Akhir tersebut.

Jember, 29 Juli 2024

Penulis



Thezar Reviansyah P.I

NIM. 2010651087

HALAMAN PERSETUJUAN SIDANG TUGAS AKHIR

**“ANALISIS SENTIMEN TERHADAP WISATA DI KABUPATEN
BANYUWANGI MENGGUNAKAN METODE *NAIVE BAYES CLASSIFIER*
(Studi Kasus : Wisata Kawah Ijen)”**

Oleh :

Thezar Reviansyah Putra Irawan

2010651087

Telah disetujui bahwa laporan Tugas Akhir ini untuk diajukan pada sidang Tugas Akhir sebagai salah satu syarat kelulusan dan mendapatkan gelar

Sarjana Komputer (S.Kom)

Universitas Muhammadiyah Jember

Disetujui oleh

Pembimbing 1

Pembimbing 2

Henny Wahyu Sulistyvo, S.Kom,

M.Kom

NIDN : 0718088309

Nur Qodariyah Fitriyah, S.T., M.Kom

NIDN : 0727097501

HALAMAN PERSETUJUAN SIDANG TUGAS AKHIR

"ANALISIS SENTIMEN TERHADAP WISATA DI KABUPATEN
BANYUWANGI MENGGUNAKAN METODE *NAIVE BAYES*
CLASSIFIER (Studi Kasus : Wisata Kawah Ijen)"

Oleh :

Thezar Reviansyah Putra Irawan

2010651087

Telah disetujui bahwa laporan Tugas Akhir ini untuk diajukan pada sidang Tugas Akhir sebagai salah satu syarat kelulusan dan mendapatkan gelar

Sarjana Komputer (S.Kom)

Universitas Muhammadiyah Jember

Disetujui oleh

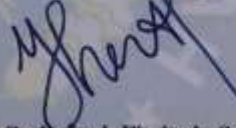
Pembimbing 1

Pembimbing 2



Henny Wahyu Sulistyo, S.Kom,
M.Kom

NIDN : 0718088309



Nur Odariyah Fitriyah, S.T., M.Kom

NIDN : 0727097501

LEMBAR PENGESAHAN
“ANALISIS SENTIMEN TERHADAP WISATA DI KABUPATEN
BANYUWANGI MENGGUNAKAN METODE *NAIVE BAYES CLASSIFIER*
(Studi Kasus : Wisata Kawah Ijen)”

Oleh :

Thezar Reviansyah Putra Irawan

2010651087

Telah disetujui bahwa laporan Tugas Akhir ini untuk diajukan pada sidang Tugas
Akhir sebagai salah satu syarat kelulusan dan mendapatkan gelar

Sarjana Komputer (S.Kom)

Universitas Muhammadiyah Jember.

Pembimbing 1

Pembimbing 2

Henny Wahyu Sulistyvo, S.Kom,
M.Kom

NIDN : 0718088309

Nur Qodariyah Fitriyah, S.T.,
M.Kom

NIDN : 0727097501

Penguji 1

Penguji 2

Guruh Wijaya, S.T., M.Kom
NIDN : 0729017601

Habibatul Azizah Al Faruq, M.Pd
NIDN : 0718128901

Dekan Fakultas Teknik

Kaprodi Teknik Informatika

Dr. Ir. Muhtar.S.T., MT., IPM
NIDN : 0010067301

Rosita Yanuarti, S.Kom., M. Cs
NIDN : 0629018601

LEMBAR PENGESAHAN

"ANALISIS SENTIMEN TERHADAP WISATA DI KABUPATEN
BANYUWANGI MENGGUNAKAN METODE *NAIVE BAYES*
CLASSIFIER (Studi Kasus : Wisata Kawah Ijen)"

Oleh :

Thezar Reviansyah Putra Irawan

2010651087

Telah disetujui bahwa laporan Tugas Akhir ini untuk diajukan pada sidang Tugas
Akhir sebagai salah satu syarat kelulusan dan mendapatkan gelar

Sarjana Komputer (S.Kom)

Universitas Muhammadiyah Jember.

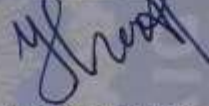
Pembimbing 1



Henny Wahyu Sulistyono, S.Kom,
M.Kom

NIDN : 0718088309

Pembimbing 2



Nur Odariyah Fitriyah, S.T.,
M.Kom

NIDN : 0727097501

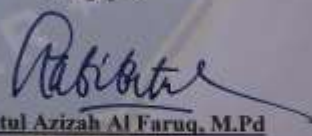
Penguji 1



Guruh Wijaya, S.T., M.Kom

NIDN : 0729017601

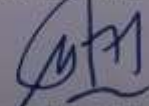
Penguji 2



Habibatul Azizah Al Faruq, M.Pd

NIDN : 0718128901

Dekan Fakultas Teknik



Dr. Ir. Muhtar, S.T., MT., IPM

NIDN : 0010067301

Kaprodi Teknik Informatika



Rosita Yauarti, S.Kom., M. Cs

NIDN : 0629018601

MOTTO

Menjadi mahasiswa adalah sebuah privilege yang tak semua pemuda diberkahi kesempatan.

(Najwa shihab)

Sholat Shubuh Jamaah nak, tidak perlu ditanya, setengah tahun aja rasakan, kalo kamu banyak gagalnya dari pada suksesnya, cacimaki saya siap.

(KH.hamid Pasuruan)

Jangan pernah khawatir dengan masa depanmu, tetap ikuti alur dan selalu libatkan allah di setiap prosesmu, Sebenarnya persaingan kita yang nyata itu ada di atas,saat semua do'a – do'a kita dipertemukan.

(Penulis)



**“ANALISIS SENTIMEN TERHADAP WISATA DI KABUPATEN
BANYUWANGI MENGGUNAKAN METODE *NAIVE BAYES CLASSIFIER*
(Studi Kasus : Wisata Kawah Ijen)”**

Thezar Reviansyah Putra Irawan¹, Henny Wahyu Sulisty², Nur Qodariyah Fitriyah³

Program Studi Teknik Informatika, Fakultas Teknik

Universitas Muhammadiyah Jember

Email : thezarreviansyah@gmail.com

ABSTRACT

Kawah Ijen Merupakan salah satu wisata yang berada di Kabupaten Banyuwangi di Provinsi Jawa Timur yang memiliki daya tarik bagi wisatawan lokal maupun mancanegara. Banyak nya komentar masyarakat atau pengunjung dari luar daerah terhadap tempat wisata Kawah Ijen dapat di analisa dengan analisis sentimen. Dalam menganalisis sentimen, penulis menggunakan algoritma *Naïve Bayes Classifier* dengan ekstraksi fitur TF-IDF. Penulis menggunakan data komentar dari Media Sosial *TripAdvisor* untuk memilih objek wisata Kawah Ijen dengan 1285 data komentar. Data tersebut akan diklasifikasi dengan label positif, negatif, dan netral menggunakan NLP(*Neuro-linguistic programming*) dan Ahli pakar. Tujuan penelitian ini adalah untuk menganalisa kecenderungan sentimen Masyarakat dan pengunjung terhadap objek Kawah Ijen. Dari hasil penelitian ini, didapatkan hasil bahwa sentimen Masyarakat serta pengunjung cenderung positive, yaitu sebesar 88,3% dengan seleksi fitur TF-IDF.

Kata Kunci : *Kawah Ijen, Analisis Sentimen, Naive Bayes Classifier, Tripadvisor*

**"SENTIMENT ANALYSIS OF TOURISM IN BANYUWANGI DISTRICT
USING THE NAIVE BAYES CLASSIFIER METHOD (Case Study: Ijen
Crater Tourism)"**

Thezar Reviansyah Putra Irawan¹, Henny Wahyu Sulisty², Nur Qodariyah Fitriyah³

Program Studi Teknik Informatika, Fakultas Teknik

Universitas Muhammadiyah Jember

Email : thezarreviansyah@gmail.com

ABSTRACT

Ijen Crater is one of the tourist attractions in Banyuwangi Regency in East Java Province which has an attraction for local and foreign tourists. Many comments from the public or visitors from outside the region regarding the Ijen Crater tourist attraction can be analyzed using sentiment analysis. In analyzing sentiment, the author uses the Naïve Bayes Classifier algorithm with TF-IDF feature extraction. The author used comment data from TripAdvisor social media to select the Ijen Crater tourist attraction with 1285 comment data. The data will be classified with positive, negative and neutral labels using NLP (Neuro-linguistic programming) and experts. The aim of this research is to analyze trends in public and visitor sentiment towards the Ijen Crater object. From the results of this research, it was found that public and visitor sentiment tended to be positive, namely 88,3% with the TF-IDF feature selection.

Keywords: *Ijen Crater, Sentiment Analysis, Naive Bayes Classifier, Tripadvisor*

KATA PENGANTAR

Assalaamu'alaikum Warahmatullahi Wabarakatuh

Puji syukur kehadirat Allah SWT yang telah melimpahkan rahmatnya kepada penulis sehingga Tugas Akhir yang berjudul “ANALISIS SENTIMEN TERHADAP WISATA DI BANYUWANGI MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER” ini dapat dilaksanakan dengan baik dan sesuai harapan. Shalawat dan salam tak lupa saya haturkan kepada junjungan kita Nabi Muhammad SAW.

Tugas Akhir ini dapat diselesaikan dengan bantuan oleh banyak pihak yang mau membantu penulis, baik secara fisik dan emosional. Oleh karena itu, ijinakan penulis berterimakasih kepada yang terhormat :

1. Bapak Dr. Ir. Muhtar.S.T., MT., IPM selaku Dekan Fakultas Teknik Universitas Muhammadiyah Jember.
2. Ibu Rosita Yanuarti, S.Kom., M. Cs selaku kepala Program Studi teknik Informatika Universitas Muhammadiyah Jember.
3. Bapak Henny Wahyu Sulisty, S.Kom, M.Kom dan Ibu Nur Qodariyah Fitriyah, S.T.,M.Kom selaku dosen pembimbing yang telah banyak berkontribusi memberikan ide, dukungan, dan masukan.
4. Bapak Guruh Wijaya, S.T., M.Kom dan Ibu Habibatul Azizah Al Faruq, M.Pd selaku dosen penguji yang telah banyak memberi masukan dan pengajaran demi keberhasilan penelitian.
5. Kedua orang tua saya terutama mama yang selalu mendoakan saya dan terus mendukung saya serta ayah dan adik, sehingga saya bisa semangat mengerjakan skripsi ini.
6. Chelsilia Melani Putri yang telah menerima keluhan kesah saya dan selalu memberikan support.
7. Teman - Teman seperjuangan saya yaitu Frangky, Riko, Renji, Ivan, Fahmi, Mirza, Hasan, Robby, Rizal Fahmi, Agil, Yoshi, dan teman – teman HIMASPR, yang selalu berada di sisi saya untuk membantu secara fisik dan emosional pada penelitian ini.

8. Serta pihak yang secara langsung maupun tidak langsung yang telah berkontribusi terhadap kelancaran tugas akhir ini.

Tugas Akhir ini pasti tidak luput dari kesalahan, oleh karena itu, penulis mengharapkan kesediaan pembaca dalam memberi masukan agar tugas akhir ini dapat menjadi lebih baik kedepannya. Semoga tugas akhir ini dapat menjadi inspirasi dan bermanfaat bagi pembaca sekalian.

Terimakasih, wassalaamu'alaikum warahmatullahi wabarakatuh.

Jember, 29 Juli 2024

Penulis

Thezar Reviansyah P.I

NIM. 2010651087



DAFTAR ISI

SURAT PERNYATAAN	i
HALAMAN PERSETUJUAN SIDANG TUGAS AKHIR	ii
DAFTAR TABEL	iii
LEMBAR PENGESAHAN	iv
MOTTO	v
ABSTRACT	vi
KATA PENGANTAR	vii
DAFTAR ISI	viii
DAFTAR GAMBAR	xi
DAFTAR TABEL	xii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan Masalah	2
1.4 Batasan Penelitian	2
1.5 Manfaat Penelitian	3
BAB II TINJAUAN PUSTAKA	4
2.1 <i>Sentimen Analysis</i>	6
2.2 Kawah Ijen	7
2.3 <i>Tripadvisor</i>	7
2.4 Data Mining	8
2.4.1 Tipe Data Mining	9
2.4.2 Manfaat Data Mining	9
2.5 Pengelompokan Data Mining	10
2.6 <i>Crawling Data</i>	11
2.7 <i>Text Mining</i>	11
2.8 <i>Classifications</i>	12
2.9 <i>Naive Bayes Classifications</i>	13
2.10 <i>Performance Evaluation Measure</i>	14

BAB III METODOLOGI PENELITIAN	16
3.1 Alur Penelitian	17
3.2 Studi Literatur	17
3.3 Observasi	17
3.4 Pengumpulan Data	20
3.5 <i>Text Processing</i>	22
3.5.1 <i>Cleaning</i>	23
3.5.2 <i>Tokenizing</i>	24
3.5.3 <i>Stopwords Removal</i>	25
3.5.4 <i>Steaming</i>	26
3.5.5 Ekstraksi Fitur	26
3.5.6 Klasifikasi <i>Naive Bayes</i>	26
3.5.7 Uji Model	27
BAB IV HASIL PENELITIAN	16
4.1 Pengambilan Data	17
4.2 Conceptual Model <i>Text Preprocessing</i>	17
4.3 Ekstrasi Fitur	17
4.4 Implementasi Klasifikasi <i>Naive Bayes</i>	20
4.5 Uji Model	22
4.6 Evaluasi Model	22
BAB V KESIMPULAN	16
5.1 Kesimpulan	17
5.2 Saran	17

DAFTAR GAMBAR

3.1 Diagram Alur Metodologi Penelitian.....	16
3.2 Halaman Awal Situs <i>Tripadvisor</i>	18
3.3 Proses <i>Crawling</i> data	19
3.5 Gambar Diagram <i>Text Pre processing</i>	22
4.1 Gambar Pembuatan <i>Word Vector</i>	26
4.2 Gambar Komentar Pengunjung Wisata Kawah ijen	27
4.3 Gambar Pengambilan Data Menggunakan <i>Web Scraper</i>	27
4.4 Gambar Diagram Pengambilan Data Pada Halaman Selanjutnya	28
4.5 Gambar Ilustrasi Pengambilan Data yang telah diambil.....	28
4.6 Gambar Diagram Hasil dari <i>Scraping</i> Data	28
4.7 Gambar Ilustrasi Alur <i>Text Preprocessing</i>	29
4.8 Gambar Diagram Kode Program Tahap <i>Cleaning</i>	30
4.9 Gambar Ilustrasi Kode Program Tahap <i>Tokenization</i>	31
4.10 Gambar Kode Program Tahap <i>Stemming</i>	32
4.11 Gambar Kode Program hasil <i>preprocessing</i>	32
4.12 Gambar Visualisasi Kata Terpopuler dengan <i>Wordcloud</i>	33
4.13 Gambar <i>Data Word Vector</i>	33
4.14 Gambar Deklarasi <i>Library</i>	36
4.15 Gambar Program <i>import</i> dataset	36
4.16 Gambar <i>CountVectorizer, TfidfTransformer, Multinomial NB</i>	37
4.17 Gambar Akurasi <i>Confusion Matrix</i>	38
4.18 Gambar Akurasi <i>Confusion Matrix</i>	38

DAFTAR TABEL

2.1 Penelitian Terdahulu	4
3.1 Waktu dan <i>Tools</i> Penelitian	18
3.2 Perangkat Keras dan Spesifikasi	19
3.3 Perangkat Lunak dan Spesifikasi	19
3.4 Proses <i>Labelling</i> data	21
3.5 Proses <i>Cleaning</i> data	22
3.6 Proses <i>Tokenizing</i>	24
3.7 Proses <i>Stopword Removal</i>	25
3.8 Proses <i>Stemming</i>	27
4.1 Pembuatan <i>Word Vector</i>	35
4.2 Tabel <i>TF (Term Frequency)</i>	35
4.3 Tabel <i>DF (Document Frequency)</i>	35
4.4 Tabel <i>IDF (Inverse Document Frequency)</i>	36
4.5 Contoh Hasil Perhitungan <i>TF-IDF</i>	36
4.6 <i>Word Vector</i> yang sudah terbobot	36
4.7 Tabel Hasil <i>Confusion Matrix</i>	38
4.8 Tabel <i>Confusion Matrix</i> Kelas <i>Positive</i>	39
4.9 Tabel <i>Confusion Matrix</i> Kelas <i>Negative</i>	39
4.10 Tabel <i>Confusion Matrix</i> Kelas <i>Netral</i>	39
4.11 Tabel Nilai Presisi, <i>Recall</i> , dan <i>f1-Score</i>	41