

**ANALISIS SENTIMEN TERHADAP TEMPAT WISATA GUNUNGKIDUL
DARI KOMENTAR PENGUNJUNG MENGGUNAKAN ALGORITMA
NAÏVE BAYES**

SKRIPSI



disusun oleh

Ma'ruf Bagas Anggadha

16.11.0477

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2021**

**ANALISIS SENTIMEN TERHADAP TEMPAT WISATA GUNUNGKIDUL
DARI KOMENTAR PENGUNJUNG MENGGUNAKAN ALGORITMA
NAÏVE BAYES**

SKRIPSI



disusun oleh

Ma'ruf Bagas Anggadha

16.11.0477

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2021**

PERSETUJUAN

SKRIPSI

**ANALISIS SENTIMEN TERHADAP TEMPAT WISATA GUNUNGKIDUL
DARI KOMENTAR PENGUNJUNG MENGGUNAKAN ALGORITMA
NAÏVE BAYES**

yang dipersiapkan dan disusun oleh

Ma'ruf Bagas Anggadha

16.11.0477

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 23 Feb 2021

Dosen Pembimbing,

Ike Verawati, M.Kom.

NIK. 19002237

PENGESAHAN

SKRIPSI

ANALISIS SENTIMEN TERHADAP TEMPAT WISATA GUNUNGGKIDUL DARI KOMENTAR PENGUNJUNG MENGUNAKAN ALGORITMA NAÏVE BAYES

Yang dipersiapkan dan disusun oleh

Ma'ruf Bagas Anggadha

16.11.0477

telah dipertahankan di depan Dewan Penguji
pada tanggal 23 Februari 2021

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Irma Rofni Wulandari, S.Pd., M.Eng
NIK.190302329

Anna Baita, M.Kom
NIK.190302290

Ike Verawati, M.Kom
NIK.19002237

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 23 Februari 2021

DEKAN FAKULTAS ILMU KOMPUTER

Windha Mega Pradnya Duhita, M.Kom
NIK. 190302185

PERNYATAAN

Saya yang bertanda tangan di bawah ini menyatakan, skripsi ini merupakan karya saya sendiri (ASLI) dan isi didalam skripsi tidak terdapat karya yang pernah diajukan oleh orang lain untuk memperoleh gelar akademis di suatu Instansi Pendidikan dan sepanjang pengetahuan saya tidak terdapat karya atau pendapat yang pernah ditulis / diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Segala sesuatu yang terkait dengan naskah dan karya yang telah dibuat adalah menjadi tanggung jawab saya pribadi.

Yogyakarta, 16 Juni 2022



Ma'ruf Bagas Anggadha

NIM 16.11.0477

MOTTO

“Jadilah Penikmat Proses, Tanpa Harus Protes”

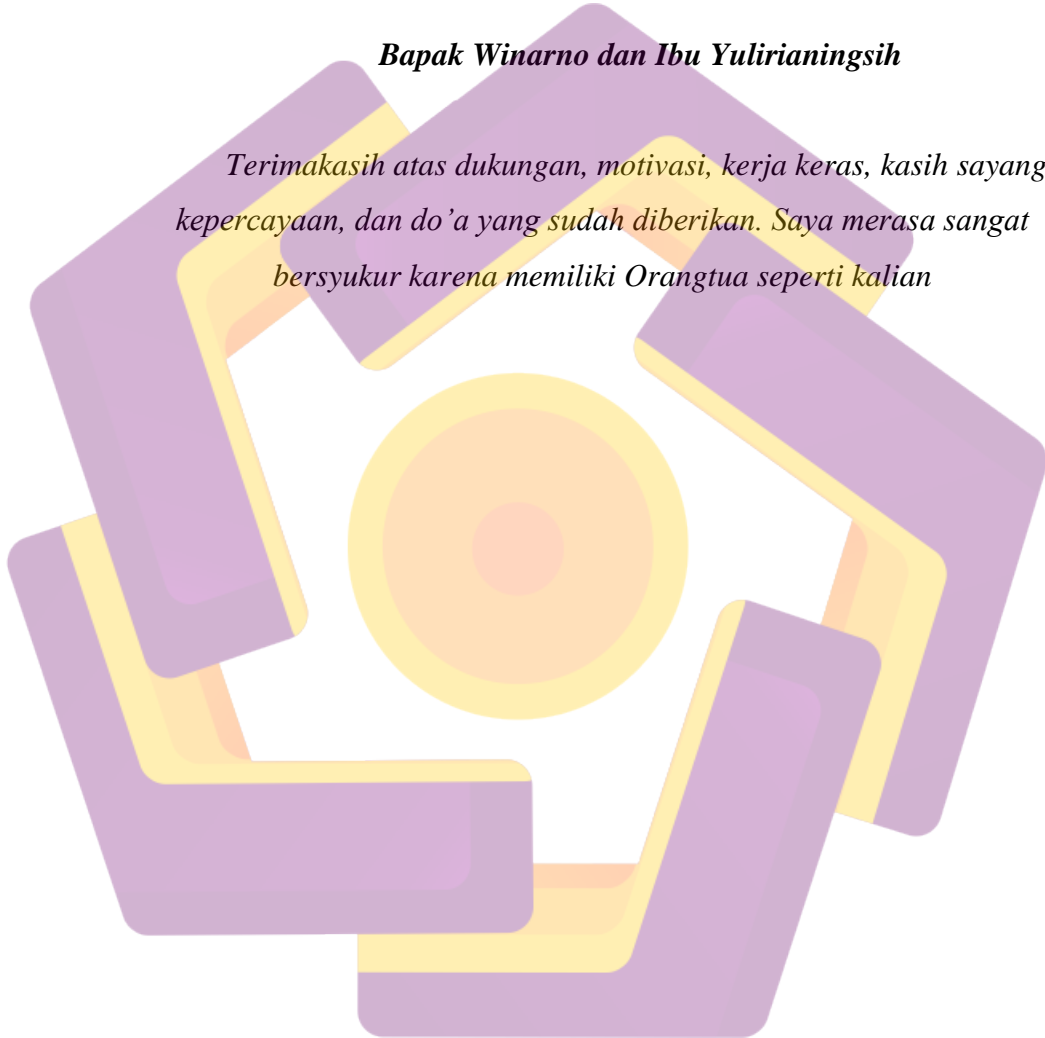


PERSEMBAHAN

Tak henti – hentinya saya mengucapkan syukur kepada Allah SWT yang telah memberikan saya kenikmatan, kesehatan, kesempurnaan, serta dapat menyelesaikan skripsi ini. skripsi ini saya persembahkan untuk:

Bapak Winarno dan Ibu Yulirianingsih

Terimakasih atas dukungan, motivasi, kerja keras, kasih sayang, kepercayaan, dan do'a yang sudah diberikan. Saya merasa sangat bersyukur karena memiliki Orangtua seperti kalian



KATA PENGANTAR

Puji dan syukur penulis panjatkan kepada Allah SWT yang telah memberikan nikmat iman, islam, dan kesehatan sehingga penulis dapat menyelesaikan skripsi ini tanpa kendala apapun. Tidak lupa sholawat serta salam penulis hanturkan kepada junjungan umat yaitu Nabi Muhammad SAW, yang telah membawa umatnya dari jaman kegelapan kepada jaman yang terang yaitu jalan yang diridhoi oleh Allah SWT.

Skripsi ini disusun sebagai salah satu syarat kelulusan bagi setiap mahasiswa Universitas Amikom Yogyakarta. Selain itu juga merupakan suatu bukti bahwa mahasiswa telah menyelesaikan kuliah jenjang program Strata-1 dan untuk memperoleh gelar Sarjana Komputer.

Dengan selesainya skripsi ini, maka penulis tidak lupa mengucapkan banyak- banyak terimakasih kepada:

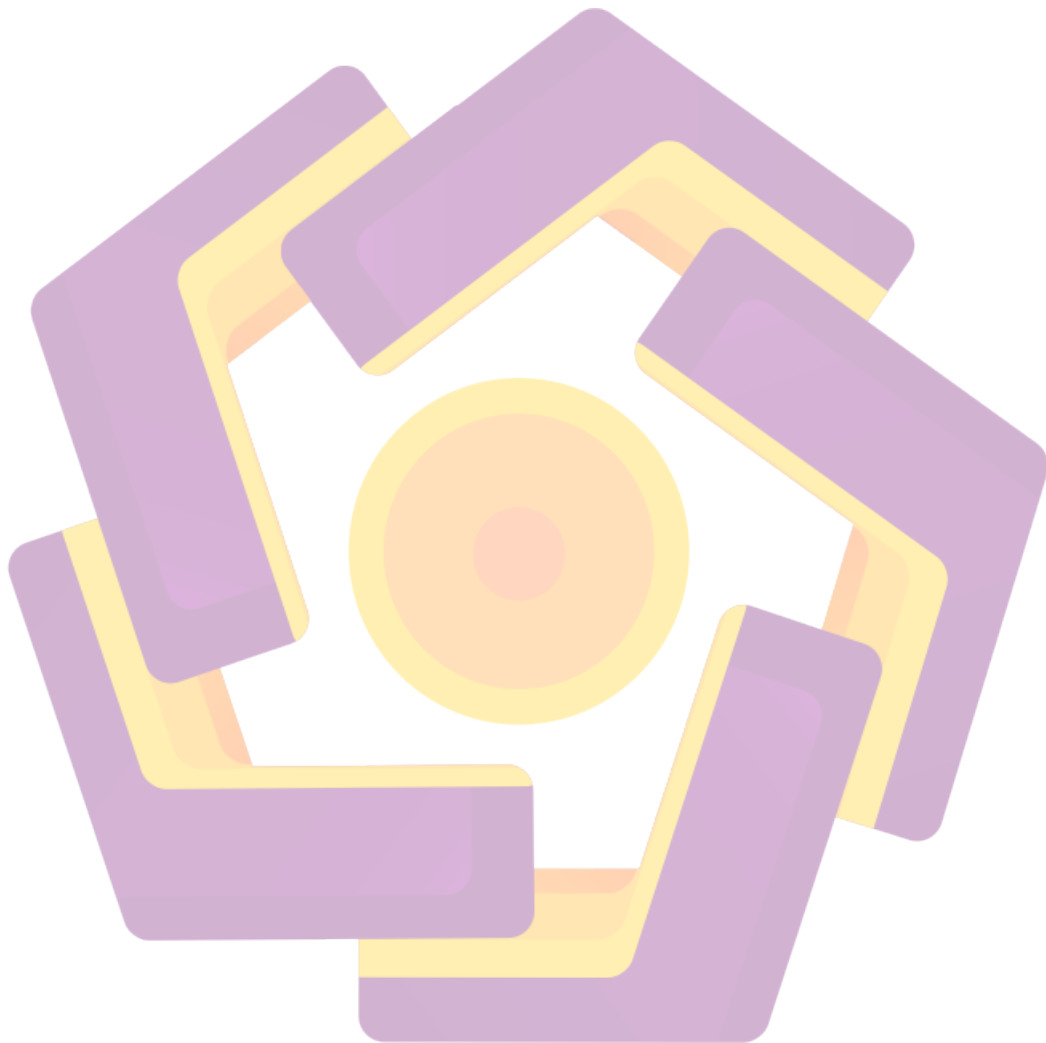
1. Bapak Prof. Dr. M. Suyanto, MM. selaku ketua Universitas Amikom Yogyakarta.
2. Ibu Ike Verawati, M. Kom selaku dosen pembimbing.
3. Bapak Winarno dan Ibu Yulirianingsih yang selalu memberikan do'a dan dukungan kepada penulis.
4. Bapak dan Ibu Dosen Universitas Amikom Yogyakarta yang telah banyak memberikan ilmu selama perkuliahan.
5. Teman-teman 16 Informatika 08 yang selalu memberikan dukungan selama proses pengerjaan.

DAFTAR ISI

PERSETUJUAN	iii
PENGESAHAN	iv
PERNYATAAN	v
MOTTO	vi
PERSEMBAHAN	vii
KATA PENGANTAR	viii
DAFTAR ISI	ix
DAFTAR TABEL	xii
DAFTAR GAMBAR	xiii
INTISARI	xiv
ABSTRACT	xv
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah	3
1.4 Maksud dan Tujuan Penelitian	3
1.5 Metodologi Penelitian	4
1.5.1 Metode Pengumpulan Data	4
1.5.1.1 Studi Literatur	4
1.5.1.2 Kuesioner	4
1.5.2 Metode Analisis	4
1.5.3 Metode Implementasi.....	5
1.5.4 Pengujian.....	5
1.6 Sistematika Penulisan	5
BAB II LANDASAN TEORI	7
2.1 Tinjauan Pustaka	7
2.2 Pengertian Data Mining.....	10
2.3 Pengertian Analisis Sentimen.....	10
2.4 Text Mining	11
2.5 Preprocessing.....	12
2.6 Term Weighting TF-IDF	13

2.6.1 Term Frequency (TF).....	13
2.6.2 Inverse Document Frequency (IDF)	14
2.7 Naïve Bayes Classifier.....	16
2.8 Confusion Matrix.....	17
2.9 Flowchart.....	18
2.10 Bahasa R.....	20
BAB III ANALISIS DAN PERANCANGAN.....	22
3.1 Alur Penelitian.....	22
3.2 Analisis Kebutuhan	23
3.2.1 Analisis Kebutuhan Fungsional	23
3.2.2 Analisis Kebutuhan Non-Fungsional	23
3.3 Perancangan Sistem.....	24
3.3.1 Pengumpulan Data Tweet	24
3.3.2 Diagram Alur Preprocessing	25
3.3.3 Perancangan Pembobotan TF-IDF	26
3.3.4 Diagram Alur Naïve Bayes Classifier.....	26
BAB IV IMPLEMETASI DAN PEMBAHASAN.....	29
4.1 Perancangan Sistem.....	29
4.1.1 Pengumpulan Data	29
4.1.2 Cleaning	29
4.1.3 Case Folding	31
4.1.4 Tokenizaation.....	31
4.1.5 Normalisasi	32
4.1.6 Filtering	34
4.1.7 Steaming.....	35
4.1.8 TF-IDF	36
4.1.9 Naïve Bayes Classifier	46
4.1.10 Pengujian Model Klasifikasi.....	56
4.1.11 Evaluasi Hasil Confusion Matrix.....	57
4.2 <i>Interface</i>	58
4.2.1 WordCloud.....	58
4.2.2 Setimen Analisis	60
4.2.3 Source Code	63
BAB V KESIMPULAN DAN SARAN.....	67

5.1 Kesimpulan.....	67
5.2 Saran.....	67
DAFTAR PUSTAKA.....	69

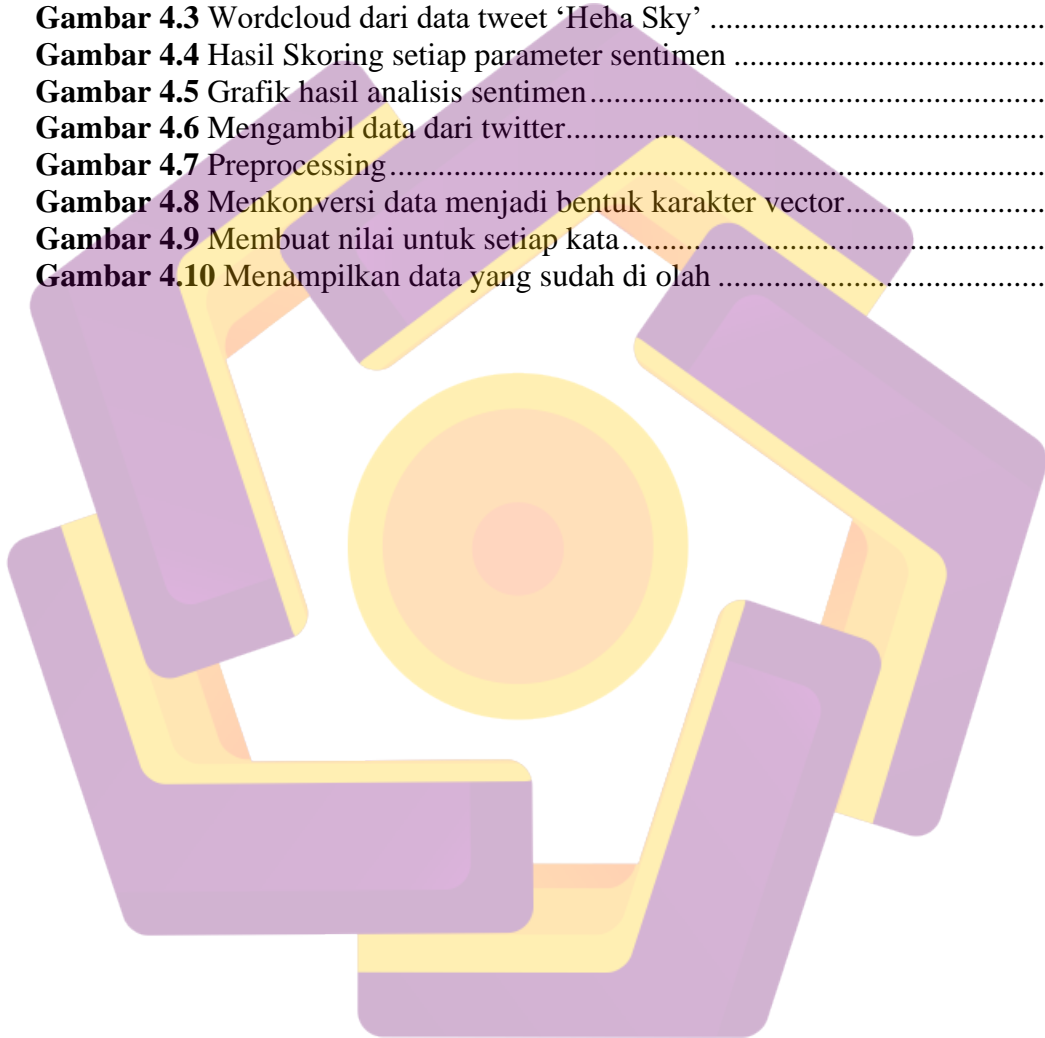


DAFTAR TABEL

Tabel 1.1 Kuesioner Pendapat Masyarakat Terhadap Wisata Yang Berada Di Kabupaten Gunungkidul	2
Tabel 2.1 Tabel Perbaningan	9
Tabel 2.2 Daftar Simbol Flowchart	18
Tabel 3.1 Data Tweet	24
Tabel 4.1 Data Tweet Sebelum Dilakukan Cleaning Data.....	30
Tabel 4.2 Data Tweet Setelah Dilakukan Cleaning Data	30
Tabel 4.3 Data Tweet Setelah Dilakukan Case Folding.....	31
Tabel 4.4 Data Tweet Setelah Dilakukan Tokenization.....	32
Tabel 4.5 Data Tweet Setelah Dilakukan Normalisasi.....	33
Tabel 4.6 Data Tweet Setelah Dilakukan Filtering	34
Tabel 4.7 Data Tweet Setelah Dilakukan Steaming	35
Tabel 4.8 Data Vector Feature.....	36
Tabel 4.9 Term Frequency Positif	37
Tabel 4.10 Term Frequency Negatif	38
Tabel 4.11 Data Document Frequency	39
Tabel 4.12 Hasil Perhitungan IDF	40
Tabel 4.13 TF-IDF Positif	42
Tabel 4.14 TF-IDF Negatif	44
Tabel 4.15 Data Uji	56
Tabel 4.16 Adalah Hasil Klasifikasi.....	56
Tabel 4.17 Confusion Matrix	57
Tabel 4.18 Hasil Precision, Recall	57

DAFTAR GAMBAR

Gambar 3.1 Alur Penelitian	22
Gambar 3.2 Alur Preprocessing.....	25
Gambar 3.3 Alur Klasifikasi Naive Bayes	27
Gambar 4.1 Data yang sudah di bersihkan	58
Gambar 4.2 Kata yang sering muncul	59
Gambar 4.3 Wordcloud dari data tweet 'Heha Sky'	60
Gambar 4.4 Hasil Skoring setiap parameter sentimen	61
Gambar 4.5 Grafik hasil analisis sentimen	62
Gambar 4.6 Mengambil data dari twitter.....	63
Gambar 4.7 Preprocessing.....	64
Gambar 4.8 Menkonversi data menjadi bentuk karakter vector.....	64
Gambar 4.9 Membuat nilai untuk setiap kata.....	65
Gambar 4.10 Menampilkan data yang sudah di olah	65



INTISARI

Penelitian ini dilatar belakangi untuk memanfaatkan teknologi sosial media dalam hal mengetahui respon dan pendapat masyarakat terhadap objek pariwisata di Gunungkidul. Salah satu informasi yang dapat digunakan adalah informasi dari media sosial. Twitter adalah raksasa utama media sosial, dengan jutaan pengguna dan memberikan wawasan yang unik, data yang ada di dalam Twitter bisa membantu mengetahui sentimen pasar secara objektif. Adapun tujuan dari penelitian ini adalah untuk menganalisis sentimen dari pengguna Twitter menggunakan algoritma naïve bayes, dan membantu calon wisatawan yang hendak berkunjung ke objek pariwisata di Gunungkidul dengan menganalisa sentimen dan menghasilkan polaritas berbentuk Grafik dan Wordcloud.

Pada penelitian ini menggunakan metode penelitian kualitatif dengan metode pengumpulan data yaitu studi literatur, dan kuesioner. Bahasa pemrograman yang digunakan adalah bahasa pemrograman R.

Hasil dari penelitian ini adalah sebuah sistem berbasis web yang dapat menampilkan polaritas dari tweet apakah bernilai positif ataupun negatif yang dapat memberikan kemudahan dalam mengetahui pendapat pengguna Twitter mengenai kata kunci atau topik yang dimasukkan.

Kata Kunci: Analisis Sentimen, *Naïve Bayes Classifier*, *Preprocessing*, *TF-IDF*, *Text mining*, *Confusion Matrix*

ABSTRACT

This research is motivated to utilize social media technology in terms of knowing the response and public opinion on tourism objects in Gunungkidul. One of the information that can be used is information from social media. Twitter is a major social media giant, with millions of users and providing unique insights, the data that is inside Twitter can help determine market sentiment objectively. The purpose of this study is to analyze the sentiment of Twitter users using the naïve Bayes algorithm, and to help potential tourists who want to visit tourism objects in Gunungkidul by analyzing sentiment and producing polarity for Graphics and Wordcloud.

In this study, using qualitative research methods with data collection methods, namely literature studies, and questionnaires. The programming language used is programming language R.

The result of this research is a web-based system that can display the polarity of a tweet whether it is positive or negative which can make it easier to find out what Twitter users think about the keywords or topics that are entered.

Keywords: *Sentiment Analysis, Naïve Bayes Classifier, Preprocessing, TF-IDF, Text mining, Confusion Matrix*

