

**ANALISIS SENTIMEN DESTINASI WISATA KULINER YOGYAKARTA
MENGUNAKAN MULTINOMIAL NAÏVE BAYES**



Diajukan sebagai salah satu syarat untuk memperoleh
Gelar Sarjana Komputer (S.Kom.)

Muhammad Aqsa Hamdani

00000032922

UMMN

UNIVERSITAS

MULTIMEDIA

NUSANTARA

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA**

TANGERANG

2023

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Muhammad Aqsa Hamdani

Nomor Induk Mahasiswa : 00000032922

Program Studi : Informatika

Skripsi dengan judul:

Analisis Sentimen Destinasi Wisata Kuliner Yogyakarta Menggunakan Multinomial Naïve 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.

UMM

Tangerang, 3 Januari 2023



(Muhammad Aqsa Hamdani)

UNIVERSITAS
MULTIMEDIA
NUSANTARA

HALAMAN PENGESAHAN

Skripsi dengan judul

ANALISIS SENTIMEN DESTINASI WISATA KULINER YOGYAKARTA MENGUNAKAN MULTINOMIAL NAÏVE BAYES

oleh

Nama : Muhammad Aqsa Hamdani
NIM : 00000032922
Program Studi : Informatika
Fakultas : Fakultas Teknik dan Informatika

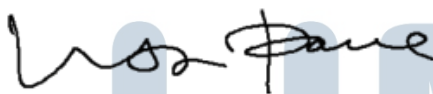
Telah diujikan pada hari Jumat, 13 Januari 2023

Pukul 13.00 s/s 15.00 dan dinyatakan

LULUS

Dengan susunan penguji sebagai berikut

Ketua Sidang



(Dr. Ivrensa Zuhdi Pane, B.Eng.,
M.Eng.)

NIDN: 08812520016

Dosen Pembimbing I

Penguji



(Suwito Pomalingo, S.Kom., M.Kom.)

NIDN: 0911098201

Dosen Pembimbing II

(Fenina Adline Twince Tobing, M.Kom.) (Marlinda Vasty Overbeek, S.Kom., M.Kom.)

NIDN: 0406058802

NIDN: 0818038501

Ketua Program Studi Informatika,

(Marlinda Vasty Overbeek, S.Kom., M.Kom.)

NIDN: 0818038501

**HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK
KEPENTINGAN AKADEMIS**

Sebagai sivitas akademik Universitas Multimedia Nusantara, saya yang bertanda tangan di bawah ini:

Nama : Muhammad Aqsa Hamdani
NIM : 00000032922
Program Studi : Informatika
Fakultas : Teknik dan Informatika
Jenis Karya : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada **Universitas Multimedia Nusantara** hak Bebas Royalti Non-eksklusif (*Non-exclusive Royalty-Free Right*) atas karya ilmiah saya yang berjudul:

**ANALISIS SENTIMEN DESTINASI WISATA KULINER YOGYAKARTA
MENGUNAKAN MULTINOMIAL NAÏVE BAYES**

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Non eksklusif ini Universitas Multimedia Nusantara berhak menyimpan, mengalih media / format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis / pencipta dan sebagai pemilik Hak Cipta. Demikian pernyataan ini saya buat dengan sebenarnya.

Tangerang, 3 Januari 2023

Yang menyatakan

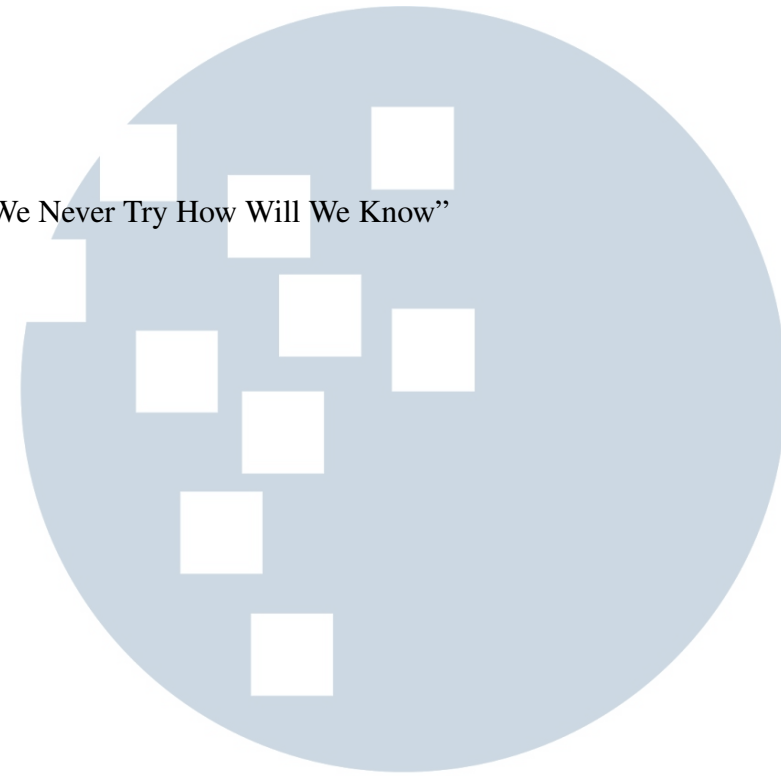


Muhammad Aqsa Hamdani

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

Halaman Persembahan / Motto

”If We Never Try How Will We Know”



UMMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA

KATA PENGANTAR

Puji Syukur atas berkat dan rahmat kepada Tuhan Yang Maha Esa, atas selesainya penulisan laporan Skripsi ini dengan judul: Analisis Sentimen Destinasi Wisata Kuliner Yogyakarta Menggunakan Multinomial Naïve Bayes dilakukan untuk memenuhi salah satu syarat untuk mencapai gelar Sarjana Komputer Jurusan Informatika Pada Fakultas Teknik dan Informatika Universitas Multimedia Nusantara. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan skripsi ini, sangatlah sulit bagi saya untuk menyelesaikan skripsi ini. Oleh karena itu, saya mengucapkan terima kasih kepada:

1. Bapak Dr. Ninok Leksono, selaku Rektor Universitas Multimedia Nusantara.
2. Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Dekan Fakultas Teknik dan Informatika Universitas Multimedia Nusantara.
3. Ibu Marlinda Vasty Overbeek, S.Kom., M.Kom., selaku Ketua Program Studi Informatika Universitas Multimedia Nusantara dan pembimbing kedua yang telah banyak membantu dan memberikan bimbingan atas terselesainya skripsi ini.
4. Ibu Fenina Adline Twince Tobing, M.Kom., sebagai Pembimbing pertama yang telah banyak meluangkan waktu untuk memberikan bimbingan, arahan dan motivasi atas terselesainya skripsi ini.
5. Kepada Diri saya Muhammad Aqsa Hamdani yang selalu *Happy Kiyowo*
6. Orang Tua tercinta saya dan keluarga saya yang selalu memberikan bantuan dukungan material dan moral, sehingga dapat menyelesaikan skripsi ini.
7. Teman-teman tercinta saya di UMN, Raisa, Dinda, islahatul, Araffi, Riyan, Danke, Zhikri, Faiz, Fathur. Terimakasih atas kebersamaan dan kekompakannya dalam melewati setiap tugas kuliah yang datang menghadang. Semoga skripsi ini bukan akhir dari perjuangan kita, sukses untuk kalian.
8. Teman-teman tercinta saya di UMN, Raisa, Dinda, islahatul, Araffi, Riyan, Danke, Zhikri, Faiz, Fathur. Terimakasih atas kebersamaan dan kekompakannya dalam melewati setiap tugas kuliah yang datang menghadang. Semoga skripsi ini bukan akhir dari perjuangan kita, sukses untuk kalian.

9. Bestie tercinta saya Ajir, Jose, dan Wulan yang selalu menjadi pendukung saya dalam mengerjakan skripsi ini.

Semoga skripsi ini berguna sebagai sumber informasi maupun sumber inspirasi, bagi para pembaca.

Tangerang, 3 Januari 2023



Muhammad Aqsa Hamdani



ANALISIS SENTIMEN DESTINASI WISATA KULINER YOGYAKARTA MENGUNAKAN MULTINOMIAL NAÏVE BAYES

Muhammad Aqsa Hamdani

ABSTRAK

Wisata kuliner Yogyakarta masih menjadi daya tarik bagi wisatawan. Media sosial Twitter menjadi wadah referensi yang digunakan dalam mempromosikan kuliner. Beragam ulasan terkait wisata kuliner di Yogyakarta yang ditulis pengguna Twitter sulit dikelola secara manual, oleh karena itu dilakukan analisis sentimen untuk mengklasifikasi teks dari ulasan pengguna Twitter tentang wisata kuliner Yogyakarta. Penelitian mengumpulkan sejumlah ulasan terhadap kuliner Yogyakarta dengan konten bahasa Indonesia. Data yang dikumpulkan kemudian dikategorikan ke dalam kelas positif, netral dan negatif. Model yang digunakan adalah *Multinomial Naïve Bayes* dengan bantuan model pembobotan kata *Term Frequency - Inverse Document Frequency* (TF-IDF) dan dilakukan oversampling menggunakan SMOTE sebagai uji coba data. Hasil akurasi terbaik yang dihasilkan oleh model *Multinomial Naïve Bayes* dengan perbandingan rasio data latih dan data uji 80:20 berada pada 67% tanpa menggunakan SMOTE, dengan nilai *precision* 67%, *recall* 77%, *f1-score* 71%.

Kata kunci: Sentimen Analisis *Multinomial Naïve bayes*, SMOTE, Kuliner Yogyakarta, TF-IDF Vectorizer



Sentiment Analysis of Yogyakarta Culinary Destinations Using Multinomial Naïve Bayes

Muhammad Aqsa Hamdani

ABSTRACT

Yogyakarta culinary tourism is still an attraction for tourists. Social media Twitter is a reference container used in promoting culinary. Various reviews related to culinary tourism in Yogyakarta written by Twitter users are difficult to manage manually, therefore a sentiment analysis is carried out to classify the text of Twitter user reviews about culinary tourism in Yogyakarta. The research collects a number of reviews of Yogyakarta culinary delights with Indonesian language content. The data collected is then categorized into positive, neutral and negative classes. The model used is *Multinomial Naïve Bayes* with the help of the word weighting model *Term Frequency - Inverse Document Frequency* (TF-IDF) and oversampling is performed using SMOTE as a data test. The best accuracy result produced by the *Multinomial Naïve Bayes* model with a ratio of training data and test data of 80:20 is at 67% without using SMOTE, with a value of *precision* 67%, *recall* 77%, *f1-score* 71%.

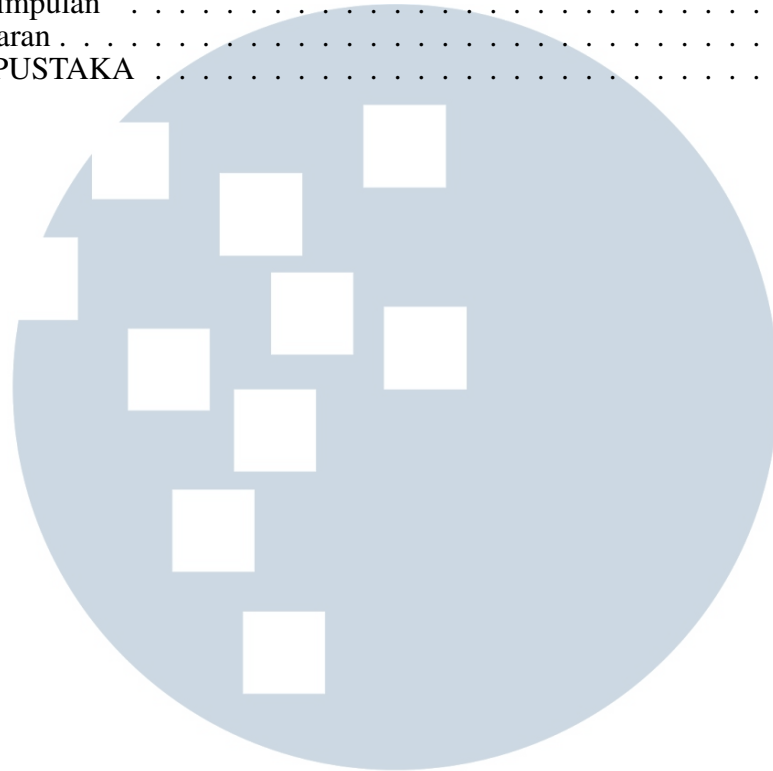
Keywords: Sentiment Analysis, Multinomial Naïve bayes, SMOTE, Yogyakarta Culinary, TF-IDF Vectorizer



DAFTAR ISI

PERNYATAAN TIDAK MELAKUKAN PLAGIAT	i
HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN	ii
HALAMAN PERSETUJUAN PUBLIKASI ILMIAH	iii
HALAMAN PERSEMBAHAN/MOTO	iv
KATA PENGANTAR	v
ABSTRAK	vii
ABSTRACT	viii
DAFTAR ISI	ix
DAFTAR GAMBAR	xi
DAFTAR TABEL	xii
DAFTAR LAMPIRAN	xiii
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah	2
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.6 Sistematika Penulisan	3
BAB 2 LANDASAN TEORI	5
2.1 Analisis Sentimen	5
2.2 Labelling Data	5
2.3 SMOTE	5
2.4 Text Preprocessing	5
2.5 Multinomial Naïve Bayes	6
2.6 Term Frequency - Inverse Document Frequency (TF-IDF)	8
2.7 Confusion Matrix	9
BAB 3 METODOLOGI PENELITIAN	11
3.1 Gambaran Umum Alur Penelitian	11
3.2 Metodologi Penelitian	11
3.2.1 Studi Literatur	11
3.2.2 Perancangan Sistem	12
BAB 4 HASIL DAN DISKUSI	16
4.1 Spesifikasi Sistem	16
4.2 Metode Pengumpulan Data	16
4.3 Implementasi Sistem	17
4.3.1 Potongan kode text preprocessing	18
4.3.2 Potongan Kode Pembobotan TF-IDF Vectorizer	25
4.3.3 Potongan kode Pembagian Data Training dan Testing	26
4.3.4 SMOTE(Synthetic Minority Oversampling Technique) Im- balanced Data Tweet	26
4.3.5 Potongan Kode Multinomial Naïve Bayes dan Evaluasi Model	27
4.4 Uji Coba dan Evaluasi	27
4.4.1 Pengujian Dengan TF-IDF Perbandingan Data Train dan Test	27
4.4.2 Pengujian Menggunakan SMOTE Pada Data Train	29
4.4.3 Pengujian Menggunakan SMOTE Pada Data Test	31
4.4.4 Pengujian Menggunakan SMOTE Pada Data Train dan Test	33

4.4.5	Pengujian Prediksi Model Menggunakan Kalimat Tweet . .	35
BAB 5	SIMPULAN DAN SARAN	37
5.1	Simpulan	37
5.2	Saran	37
DAFTAR PUSTAKA	38



UMMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR GAMBAR

Gambar 3.1	Gambar Diagram Alur Penelitian	11
Gambar 3.2	Gambar Labelling Data	12
Gambar 3.3	Flowchart Text Pre-processing	13
Gambar 3.4	Flowchart menerapkan TF-IDF	14
Gambar 3.5	Flowchart menerapkan Multinomial Naïve Bayes	15
Gambar 4.1	Wordcloud Kalimat Kuliner Yogyakarta	25
Gambar 4.2	Train dan Test Plot Dataset	28
Gambar 4.3	Classification Report Perbandingan Rasio 80:20	28
Gambar 4.4	Confusion Matrix Perbandingan Rasio 80:20	29
Gambar 4.5	Train Diagram Plot Dataset Menggunakan SMOTE	30
Gambar 4.6	Classification Report Menggunakan SMOTE Pada Data Train	30
Gambar 4.7	Confusion Matrix Menggunakan SMOTE Oversampling Pada Data Train	31
Gambar 4.8	Test Diagram Plot Dataset Menggunakan SMOTE	32
Gambar 4.9	Classification Report Menggunakan SMOTE Pada Data Test	32
Gambar 4.10	Confusion Matrix Menggunakan SMOTE Oversampling Pada Data Test	33
Gambar 4.11	Train dan Test Diagram Plot Dataset Menggunakan SMOTE	34
Gambar 4.12	Classification Report Menggunakan SMOTE Pada Data Train dan Test	34
Gambar 4.13	Confusion Matrix Menggunakan SMOTE Oversampling Pada Data Train dan Test	35
Gambar 4.14	Hasil Kalimat Pada Data Predict Model	36
Gambar 4.15	Hasil Prediksi Pada Kalimat Tweet	36



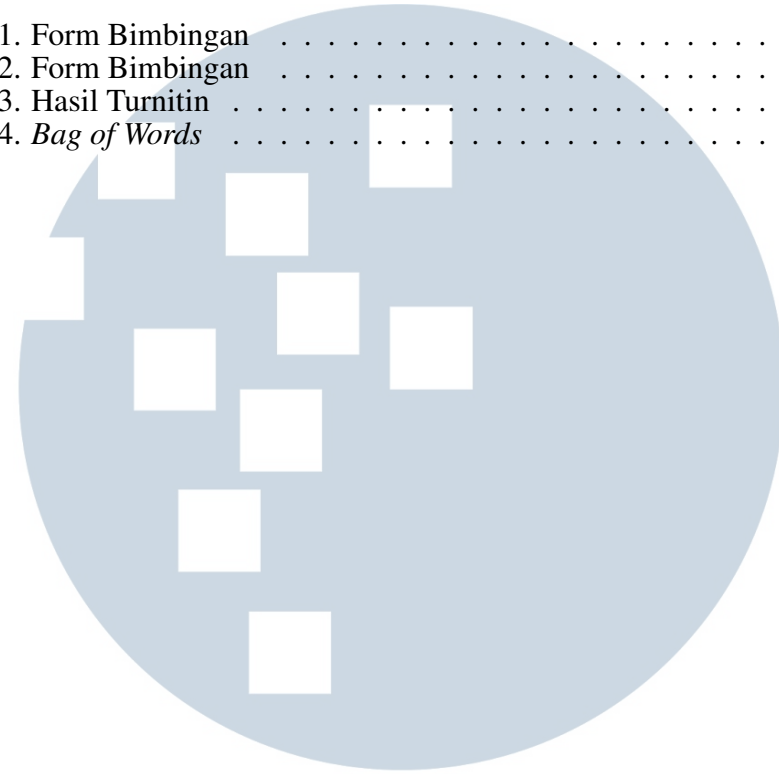
DAFTAR TABEL

Tabel 2.1	Confusion Matrix	9
Tabel 4.1	Hasil Case Folding	18
Tabel 4.2	Hasil Data Cleaning	20
Tabel 4.3	Hasil Tokenization	21
Tabel 4.4	Hasil Stopword	23
Tabel 4.5	Hasil Stemming Bahasa Indonesia	24



DAFTAR LAMPIRAN

Lampiran 1. Form Bimbingan	38
Lampiran 2. Form Bimbingan	39
Lampiran 3. Hasil Turnitin	41
Lampiran 4. <i>Bag of Words</i>	46



UMMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA