

**ANALISIS SENTIMEN TERHADAP MUSEUM SONOBUDOYO
MENGUNAKAN MODEL *MACHINE LEARNING***

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

ANNISA BUDHI PRADANI

20.11.3474

Kepada

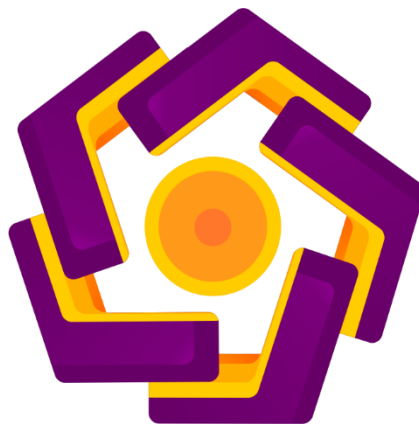
**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2024

**ANALISIS SENTIMEN TERHADAP MUSEUM SONOBUDOYO
MENGUNAKAN MODEL *MACHINE LEARNING***

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

ANNISA BUDHI PRADANI

20.11.3474

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2024

HALAMAN PERSETUJUAN

SKRIPSI

**ANALISIS SENTIMEN TERHADAP MUSEUM SONOBUDOYO
MENGUNAKAN MODEL *MACHINE LEARNING***

yang disusun dan diajukan oleh

Annisa Budhi Pradani

20.11.3474

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 24 Juli 2024

Dosen Pembimbing,



Yoga Pristiyanto, S.Nom, M.Eng
NIK. 190302412

HALAMAN PENGESAHAN

SKRIPSI

**ANALISIS SENTIMEN TERHADAP MUSEUM SONOBUDOYO
MENGUNAKAN MODEL *MACHINE LEARNING***

yang disusun dan diajukan oleh

Annisa Budhi Pradani

20.11.3474

Telah dipertahankan di depan Dewan Penguji
pada tanggal 24 Juli 2024

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Ali Mustopa, M. Kom
NIK. 190302192

Ike Verawati, M.Kom
NIK. 190302237

Yoga Pristyanto, S.Kom, M.Eng
NIK. 190302412



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 24 Juli 2024

DEKAN FAKULTAS ILMU KOMPUTER



Hanif Al Fatta, S.Kom., M.Kom., Ph.D.
NIK. 190302096

**ANALISIS SENTIMEN TERHADAP MUSEUM SONOBUDOYO
HALAMAN PERNYATAAN KEASLIAN SKRIPSI**

Yang bertandatangan di bawah ini,

Nama mahasiswa : Annisa Budhi Pradani
NIM : 20.11.3474

Menyatakan bahwa Skripsi dengan judul berikut:

Analisis Sentimen Terhadap Museum Sonobudoyo Menggunakan Model *Machine Learning*

Dosen Pembimbing : Yoga Pristyanto, S.Kom, M.Eng

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 24 Juli 2024



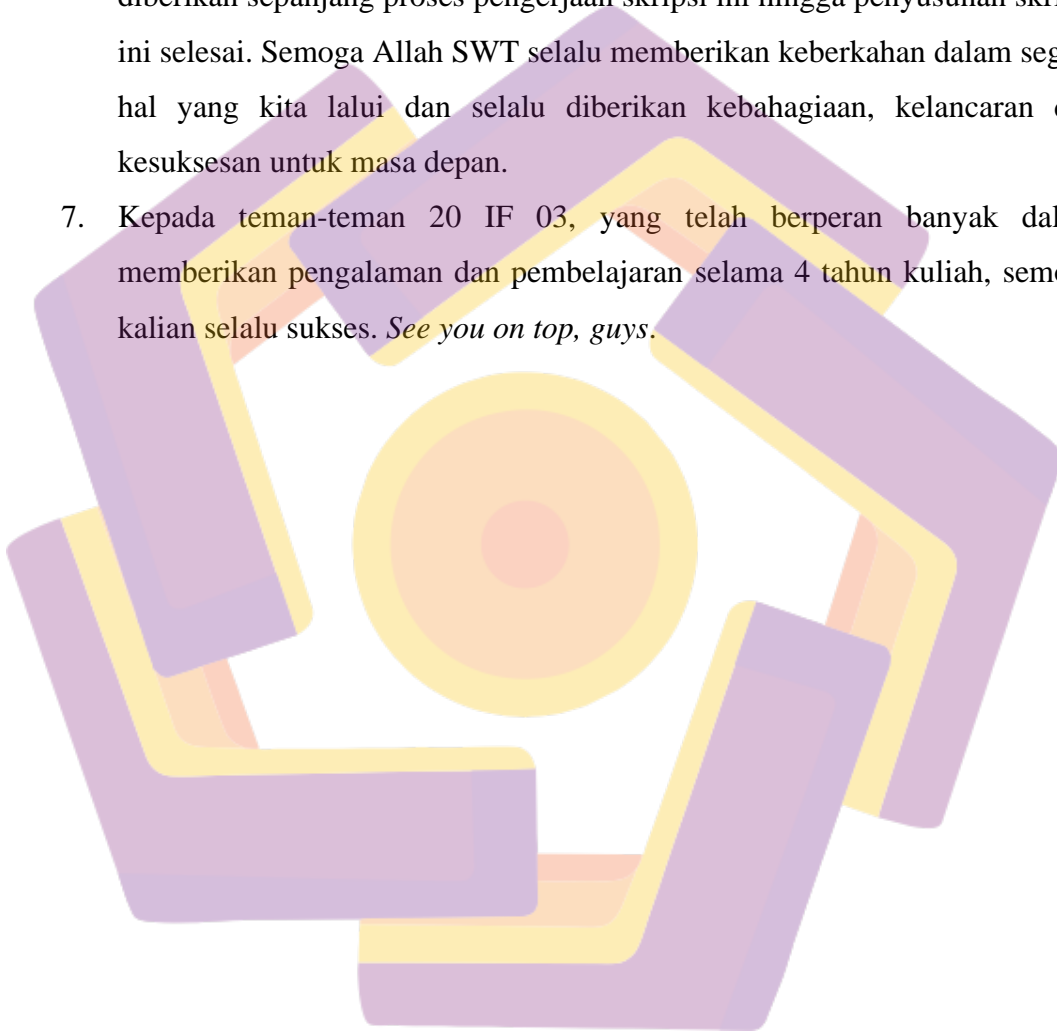
Annisa Budhi Pradani

HALAMAN PERSEMBAHAN

Puji syukur penulis panjatkan kepada Allah SWT, yang telah memberikan kesehatan, rahmat, dan hidayah, sehingga penulis masih diberikan kesempatan untuk menyelesaikan skripsi ini sebagai salah satu syarat untuk mendapatkan gelar kesarjanaan. Walaupun jauh dari kata sempurna, penulis bangga telah mencapai titik ini dan akhirnya skripsi ini bisa selesai tepat waktu. Dengan dukungan dan doa dari orang-orang tercinta, skripsi ini akhirnya selesai dengan baik dan tepat waktu. Oleh karena itu, penulis mengucapkan rasa syukur dan terima kasih kepada :

1. Kepada Tuhan Yang Maha Esa, karena atas izin dan karunia-Nya, skripsi ini dapat diselesaikan tepat waktu. Terima kasih yang tak terhingga kepada Tuhan, penguasa alam, yang telah meridhoi dan mengabulkan segala doa.
2. Kepada orang tua tersayang Bapak R. Soepono M.H dan Ibu Dra. Riharyani yang telah memberikan dukungan moril maupun materi serta doa yang tiada henti untuk kesuksesan saya. Tidak ada kata yang lebih indah dari doa, dan doa yang paling tulus datang dari orang tua. Ucapan terima kasih saja tidak cukup untuk membalas kebaikan mereka. Oleh karena itu, terimalah persembahan bakti dan cinta dari saya untuk kalian, bapak dan ibu.
3. Kepada Bapak Yoga Pristyanto, S.Kom., M.Eng selaku dosen pembimbing, terima kasih atas waktu, bimbingan, dan pelajaran yang telah diberikan kepada saya. Bimbingan dan pelajaran yang diberikan sangat berharga dan telah membantu saya menjadi lebih baik.
4. Kepada bapak dan ibu dosen Universitas Amikom Yogyakarta yang selalu tulus membantu dan memberikan ilmu selama saya menjadi mahasiswa.
5. Kepada sahabat saya Anisa Nurhayati, Oktaviana Aditya T, Siti Nadia Nurul F, dan Wissel Virginia Uktolseya. Tanpa semangat, dukungan dan bantuan kalian semua tak kan mungkin aku sampai disini, terimakasih untuk canda tawa, tangis, dan perjuangan yang kita lewati bersama dan terimakasih untuk kenangan manis yang telah mengukir selama ini. Dengan perjuangan dan kebersamaan kita pasti bisa! Semangat!.

6. Kepada orang terkasih Lintang Pamungkas Ocvianto Thiertian, yang telah menjadi salah satu penyemangat dalam suka maupun duka karena selalu menemani dan menjadi *support system* penulis yang tidak mudah selama proses pengerjaan skripsi. Terima kasih telah mendengarkan keluh kesah, menjadi orang yang sabar memberikan dukungan, waktu dan materi yang diberikan sepanjang proses pengerjaan skripsi ini hingga penyusunan skripsi ini selesai. Semoga Allah SWT selalu memberikan keberkahan dalam segala hal yang kita lalui dan selalu diberikan kebahagiaan, kelancaran dan kesuksesan untuk masa depan.
7. Kepada teman-teman 20 IF 03, yang telah berperan banyak dalam memberikan pengalaman dan pembelajaran selama 4 tahun kuliah, semoga kalian selalu sukses. *See you on top, guys.*



KATA PENGANTAR

Segala puji dan syukur kami panjatkan ke hadirat Allah SWT yang telah memberikan keimanan, rahmat, petunjuk, bimbingan, kasih sayang, nikmat, kesehatan, dan kesempatan. Kami juga mengirimkan shalawat dan salam kepada Nabi kita, Rasulullah Muhammad SAW. Berkat rahmat Allah SWT, penulis berhasil menyelesaikan skripsi yang berjudul “Analisis Sentimen Terhadap Museum Sonobudoyo Menggunakan Model Machine Learning.”

Dalam proses penyelesaian skripsi ini, penulis telah menerima berbagai bantuan, baik materi maupun moral, dari banyak pihak yang dengan tulus telah memberikan dukungan. Dengan segala kerendahan hati, penulis menyampaikan terima kasih yang sebesar-besarnya dan memohon maaf jika terdapat kekurangan dalam penulisan skripsi ini. Melalui kesempatan ini, penulis ingin mengucapkan terima kasih kepada :

1. Bapak Prof Dr. M. Suyanto, M.M selaku Rektor Universitas Amikom Yogyakarta.
2. Bapak Hanif Al Fatta, S.Kom., M.Kom selaku Dekan Fakultas Ilmu Komputer Universitas Amikom Yogyakarta
3. Ibu Windha Mega Pradnya D, M.Kom selaku ketua program studi S1 Informatika.
4. Kepada Bapak Yoga Pristyanto, S.Kom., M.Eng selaku dosen pembimbing.
5. Kepada Bapak Ery Sustiyadi, S.T., M.A. selaku kepala Museum Sonobudoyo Yogyakarta.

Yogyakarta, 24 Juli 2024

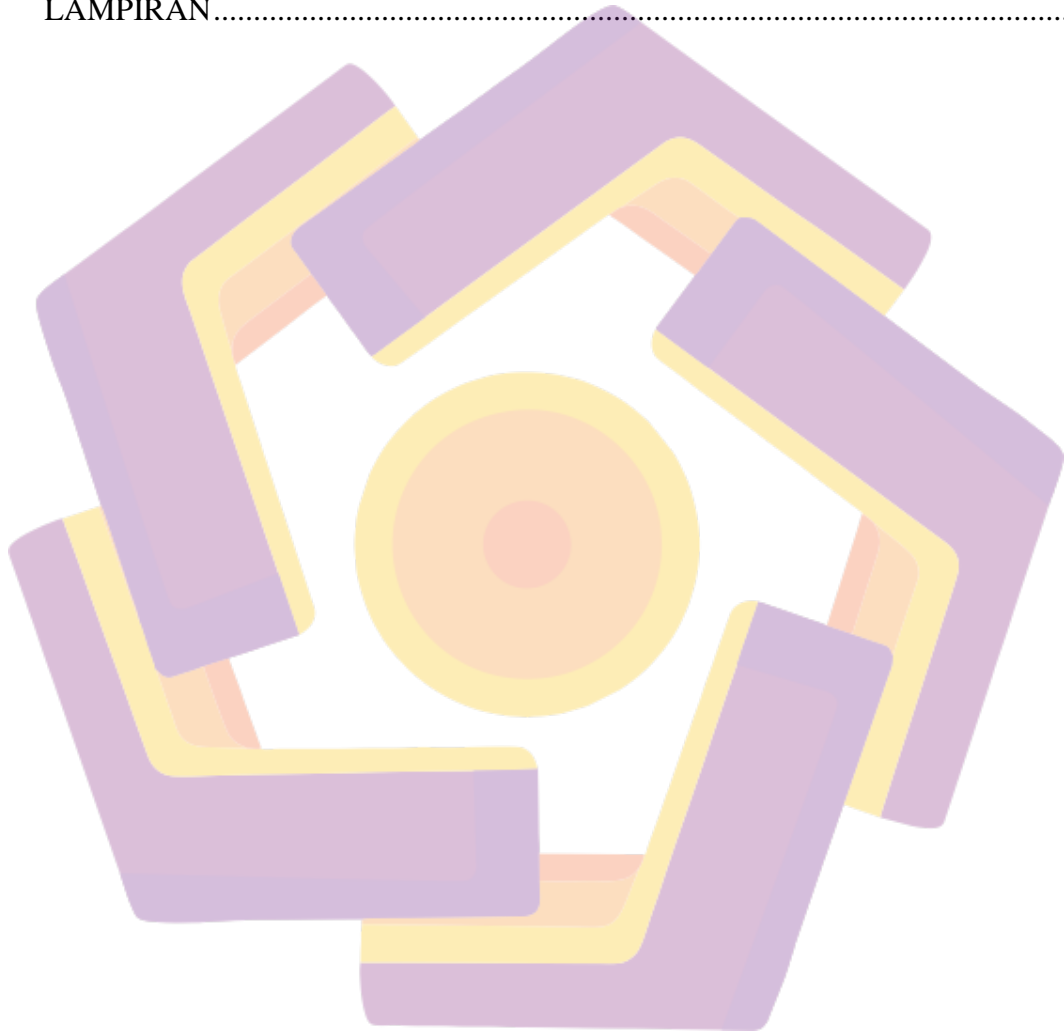
Annisa Budhi Pradani

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vii
DAFTAR ISI.....	viii
DAFTAR TABEL.....	xi
DAFTAR GAMBAR.....	xii
DAFTAR LAMPIRAN.....	xiii
DAFTAR LAMBANG DAN SINGKATAN	xiv
DAFTAR ISTILAH	xv
INTISARI	xvi
<i>ABSTRACT</i>	xvii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	2
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.6 Sistematika Penulisan	4
BAB II TINJAUAN PUSTAKA	5
2.1 Studi Literatur	5
2.2 Dasar Teori	12
2.2.1 Machine Learning	12
2.2.2 Text Mining	13
2.2.3 Analisis Sentimen	13
2.2.4 Naive Bayes Classifier	14

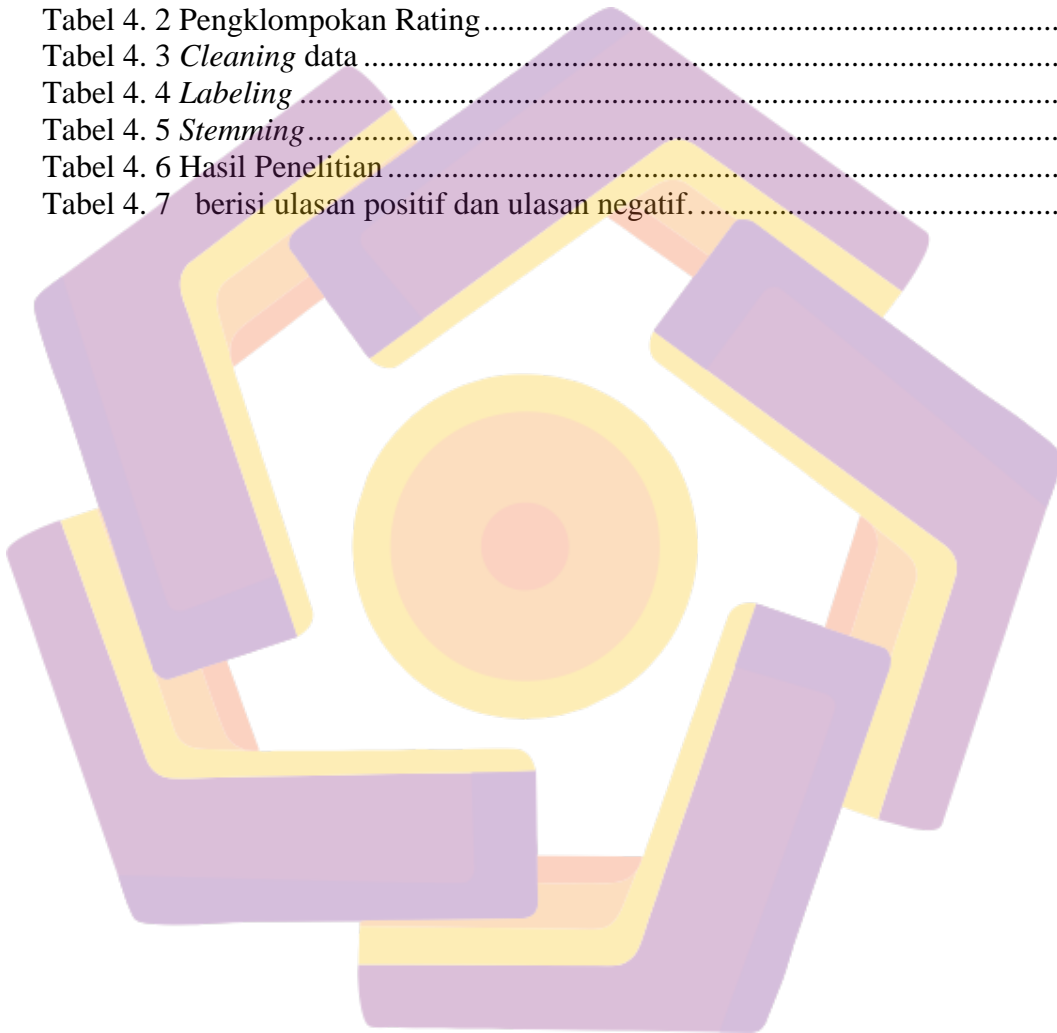
2.2.5 Model Evaluasi	15
2.2.6 TF-IDF	16
2.2.7 Web Scrapping.....	18
2.2.8 <i>Google Maps</i>	18
2.2.9 Confusion Matrix	19
2.2.10 SMOTE	20
BAB III METODE PENELITIAN	21
3.1 Objek Penelitian.....	21
3.2 Alur Penelitian	22
3.3 Alat dan Bahan.....	24
3.3.1 Alat.....	24
3.3.2 Bahan	25
BAB IV HASIL DAN PEMBAHASAN	26
4.1.1 Dataset.....	26
4.1.2 <i>Preprocessing</i>	28
4.1.2.1 <i>Cleaning</i>	30
4.1.2.2 <i>Labeling</i>	31
4.1.2.3 <i>Stemming</i>	33
4.1.3 <i>Data Prartitioning</i>	35
4.1.4 <i>Feature Extraction</i>	35
4.1.5 <i>Model Training</i>	35
4.1.5.1 MultinomialNB	36
4.1.5.2 GaussianNB	37
4.1.6 Model Evaluation.....	37
4.1.6.1 Evaluasi MultinomialNB	38
4.1.6.2 Evaluasi GaussianNB.....	39

4.1.7 Hasil Penelitian	40
BAB V PENUTUP	45
5.1 Kesimpulan	45
5.2 Saran	46
REFERENSI	47
LAMPIRAN.....	50



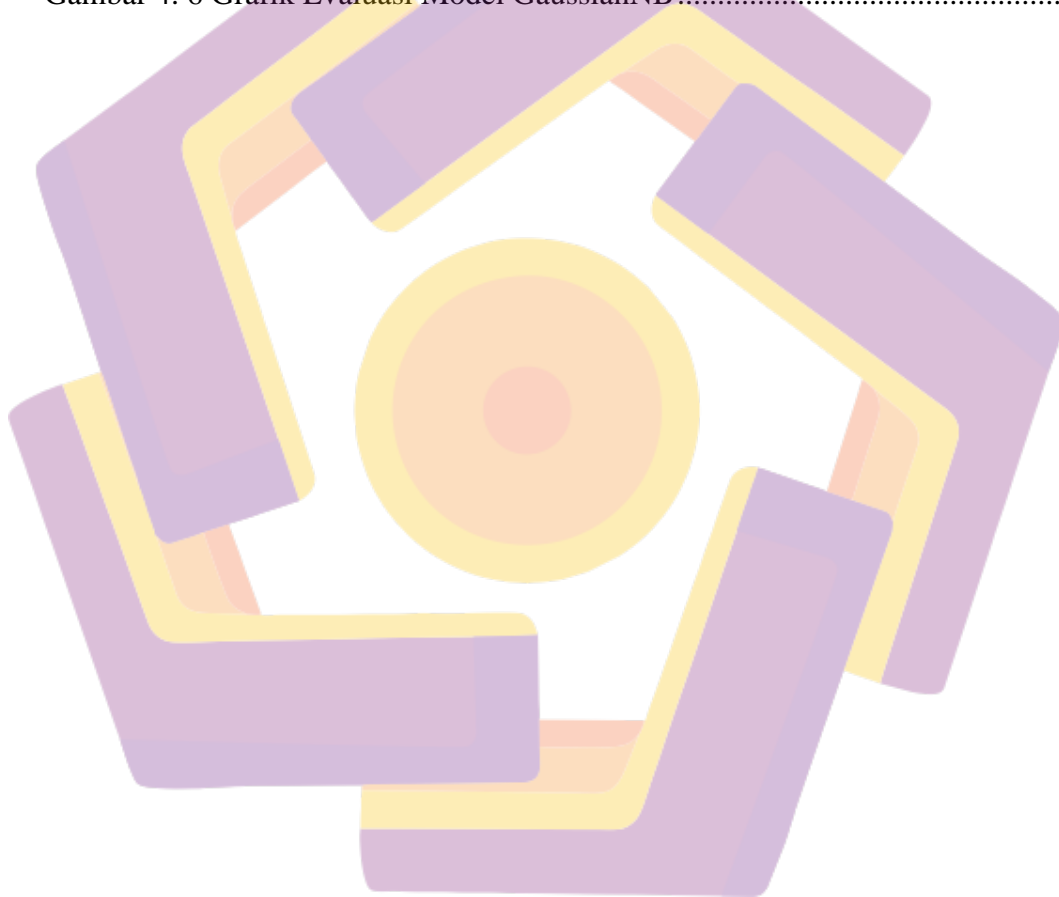
DAFTAR TABEL

Table 2.1 Tabel 2.1 Keaslian Penelitian	7
Table 2. 2 Kategori Analisis Sentimen	14
Table 2. 3 <i>Confusion Matrix</i>	19
Tabel 3. 1 Contoh ulasan pada dataset.....	25
Tabel 4. 1 Hasil <i>Scrapping</i>	27
Tabel 4. 2 Pengklompokan Rating.....	29
Tabel 4. 3 <i>Cleaning</i> data	30
Tabel 4. 4 <i>Labeling</i>	31
Tabel 4. 5 <i>Stemming</i>	33
Tabel 4. 6 Hasil Penelitian	41
Tabel 4. 7 berisi ulasan positif dan ulasan negatif.	42



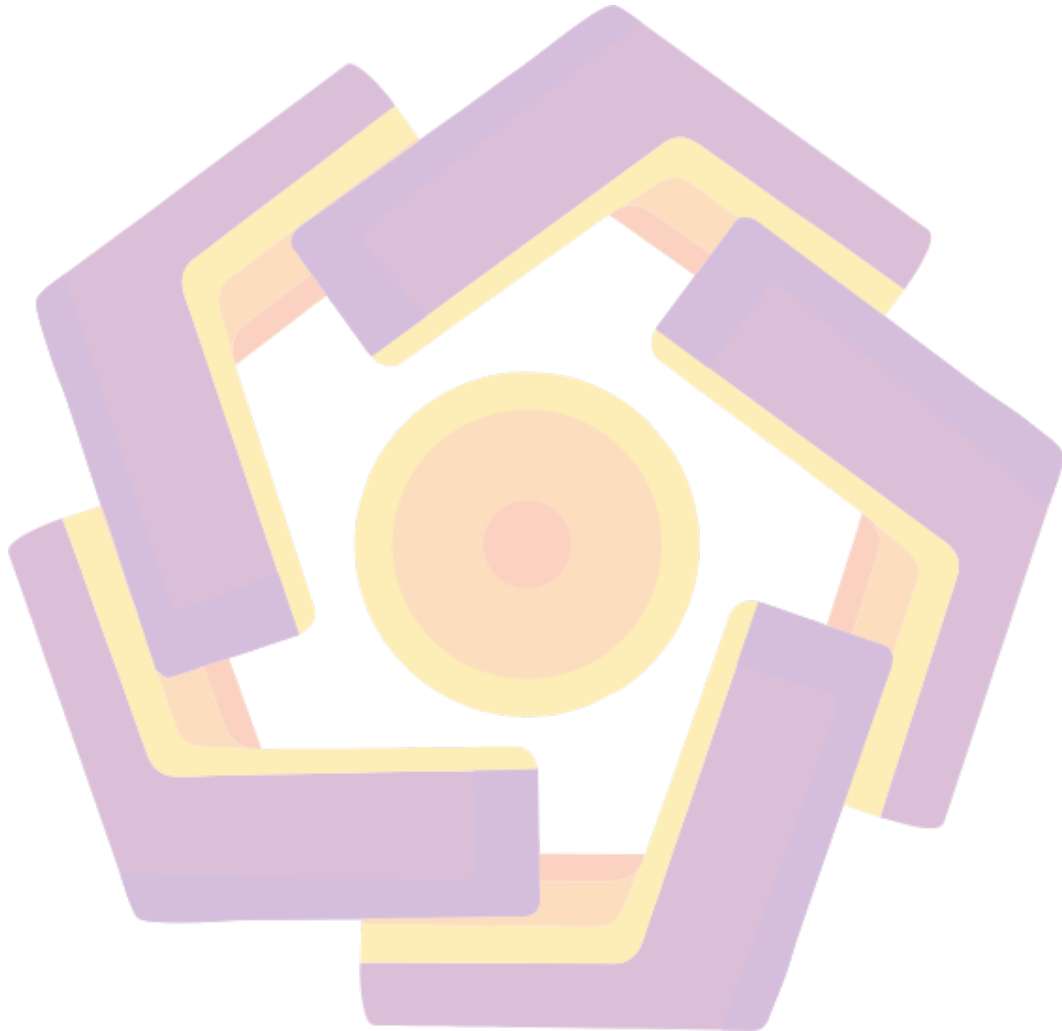
DAFTAR GAMBAR

Gambar 2. 1 Proses Web Scrapping	18
Gambar 3. 1 Alur Penelitian	22
Gambar 4. 1 Contoh <i>Review</i> Positif	27
Gambar 4. 2 Contoh <i>Review</i> Negatif	27
Gambar 4. 3 Grafik Hasil Model MultinomialNB	36
Gambar 4. 4 Grafik Hasil Model GaussianNB	37
Gambar 4. 5 Grafik Hasil Evaluasi Model MultinomialNB	39
Gambar 4. 6 Grafik Evaluasi Model GaussianNB	40



DAFTAR LAMPIRAN

Lampiran 1. Profil Objek Penelitian	50
Lampiran 2. Dokumentasi bersama kepala Museum Sonobudoyo.....	51
Lampiran 3. Dokumentasi Gedung Lama	52
Lampiran 4. Dokumentasi Gedung Baru	53
Lampiran 5. Dokumentasi Layanan dan Fasilitas	54



DAFTAR LAMBANG DAN SINGKATAN

$P(A|B)$: Peluang kejadian A apabila B terjadi, yang dikenal sebagai *posterior probability*.

$P(B|A)$: Peluang kejadian B apabila A terjadi, yang dikenal sebagai *likelihood*.

$P(B)$: Probabilitas (B), yang dikenal sebagai *prior probability*.

Berlaku ketentuan yaitu $P(B) \neq 0$.

D : Dokumen ke- d .

t : Kata (term) ke- t dalam dokumen.

W : Bobot untuk kata ke- t dalam dokumen ke- d

tf : Jumlah kemunculan kata ke- t dalam dokumen.

idf : *Inverse Document Frequency*, mengukur seberapa jarang kata ke- t muncul di seluruh dokumen.

df : Jumlah dokumen yang mengandung kata ke- t .

TP (*True Positive*) : Jumlah data nilai aktual kelas positif dan nilai prediksi kelas positif

TN (*True Negative*) : Jumlah data nilai aktual negatif dan nilai prediksi negatif

FP (*False Positive*) : Jumlah data nilai aktual positif dan nilai prediksi negatif

FN (*False Negative*) : Jumlah data nilai aktual negatif dan nilai Prediksi positif

ML : *Machine Learning*

AI : *Artificial Intelligence*

ROC-AUC : *Receiver Operating Characteristic Area Under the Curve*

MultinomialNB : Multinomial Naive Bayes

GaussianNB : Gaussian Naive Bayes

DAFTAR ISTILAH

Website	Kumpulan halaman web yang dapat diakses melalui internet.
Rating	Peringkat dari hasil penilaian yang diberikan oleh pengguna,
Review	Ulasan yang diberikan pengguna pada aplikasi.
SerpApi	Menyediakan data pencarian dari berbagai mesin pencari (Google, Bing, Yahoo).



INTISARI

Museum Sonobudoyo, didirikan pada tahun 1935, berperan penting dalam melestarikan budaya Jawa, Madura, Bali, dan Lombok. Di era digital, ulasan online menjadi sumber penting untuk menilai kepuasan pengunjung. Penelitian ini menganalisis sentimen ulasan Museum Sonobudoyo dari Google Maps untuk memahami pandangan pengunjung dan meningkatkan layanan museum. Data ulasan dikumpulkan melalui scraping dengan SerpApi dan diubah menjadi format numerik menggunakan TF-IDF, lalu dibagi menjadi data pelatihan dan pengujian. Empat model klasifikasi digunakan: Multinomial Naive Bayes (MultinomialNB), Bagged MultinomialNB, Gaussian Naive Bayes (GaussianNB), dan Bagged GaussianNB. Hasil pengujian menunjukkan bahwa Bagged MultinomialNB memiliki kinerja terbaik dengan akurasi pelatihan 96.53% dan akurasi pengujian 89.16%. Model ini juga mencatat nilai ROC-AUC 99.94% untuk pelatihan dan 93.04% untuk pengujian, serta nilai cross-validation 90.43%. Sebaliknya, Bagged GaussianNB menunjukkan kinerja yang lebih rendah. Hasil ini menunjukkan perlunya lebih banyak data sentimen negatif untuk meningkatkan kemampuan model mengenali sentimen, memberikan wawasan penting bagi pengelola Museum Sonobudoyo dalam meningkatkan layanan.

Kata kunci: Analisis Sentimen, Naive Bayes, Multinomial Naive Bayes, Gaussian Naive Bayes, Bagging, TF-IDF, Web Scraping, Google Maps, Museum Sonobudoyo

ABSTRACT

Museum Sonobudoyo, established in 1935, plays a crucial role in preserving the cultures of Java, Madura, Bali, and Lombok. In the digital era, online reviews have become an important source for assessing visitor satisfaction. This study analyzes the sentiment of Museum Sonobudoyo reviews from Google Maps to understand visitor perspectives and improve museum services. Review data were collected through scraping with SerpApi and transformed into numeric format using TF-IDF, then split into training and testing data. Four classification models were used: Multinomial Naive Bayes (MultinomialNB), Bagged MultinomialNB, Gaussian Naive Bayes (GaussianNB), and Bagged GaussianNB. The test results show that Bagged MultinomialNB performed best, with a training accuracy of 96.53% and a testing accuracy of 89.16%. This model also recorded an ROC-AUC of 99.94% for training and 93.04% for testing, along with a cross-validation score of 90.43%. In contrast, Bagged GaussianNB showed lower performance. These results indicate the need for more negative sentiment data to enhance the model's ability to recognize sentiments, providing valuable insights for Museum Sonobudoyo management to improve services.

Keyword: *Sentiment Analysis, Naive Bayes, Multinomial Naive Bayes, Gaussian Naive Bayes, Bagging, TF-IDF, Web Scraping, Google Maps, Sonobudoyo Museum*