

**ANALISIS SENTIMEN OPINI PENGUNJUNG PADA EVENT
BUDAYA POP JEPANG PADA MEDIA SOSIAL TWITTER
MENGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana Informatika



disusun oleh

AKHMAD TAAJUDDIN ZAHIR

19.11.3193

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2023

**ANALISIS SENTIMEN OPINI PENGUNJUNG PADA EVENT
BUDAYA POP JEPANG PADA MEDIA SOSIAL TWITTER
MENGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana Informatika



disusun oleh

AKHMAD TAAJUDDIN ZAHIR

19.11.3193

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2023

HALAMAN PERSETUJUAN

SKRIPSI

**ANALISIS SENTIMEN OPINI PENGUNJUNG PADA EVENT
BUDAYA POP JEPANG PADA MEDIA SOSIAL TWITTER
MENGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE**


yang disusun dan diajukan oleh

Akhmad Taajuddin Zahir

19.11.3193

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 18 Oktober 2023

Dosen Pembimbing,



Yoga Pristyanto, S.Kom, M.Eng

NIK. 190302412

HALAMAN PENGESAHAN

SKRIPSI

**ANALISIS SENTIMEN OPINI PENGUNJUNG PADA EVENT
BUDAYA POP JEPANG PADA MEDIA SOSIAL TWITTER
MENGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE**

yang disusun dan diajukan oleh

Akhmad Taajuddin Zahir

19.11.3193

Telah dipertahankan di depan Dewan Penguji
pada tanggal 18 Oktober 2023

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Anggit Dwi Hartanto, M.Kom
NIK. 190302163

Anna Baita, M. Kom
NIK. 190302290

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



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 18 Oktober 2023

DEKAN FAKULTAS ILMU KOMPUTER



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

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Akhmad Taajuddin Zahir
NIM : 19.11.3193

Menyatakan bahwa Skripsi dengan judul berikut:

Analisis Sentimen Opini Pengunjung Pada Event Budaya Pop Jepang Pada Media Sosial Twitter Menggunakan Algoritma Support Vector Machine

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, 18 Oktober 2023

Yang Menyatakan,



Akhmad Taajuddin Zahir

HALAMAN PERSEMBAHAN

Dengan rasa syukur yang mendalam atas rezeki yang melimpah yang diberikan Allah Swt sehingga dapat menyelesaikan skripsi ini. Penulis mempersembahkan skripsi ini kepada:

1. Keluarga besar penulis yang telah senantiasa memberikan doa, dukungan serta motivasi baik secara moril maupun materil selama menempuh perkuliahan ini hingga selesai.
2. Segenap *civitas* akademika kampus Universitas AMIKOM Yogyakarta, semoga selalu tetap semangat dalam beraktivitas selama berada di kampus Universitas AMIKOM Yogyakarta.
3. Teman-teman penulis baik itu seangkatan, adik kelas, kakak kelas dari Universitas AMIKOM Yogyakarta maupun teman-teman dari universitas lain yang telah memberikan semangat, masukan, dan arahan hingga akhirnya dapat terselesaikan skripsi ini.

KATA PENGANTAR

Dengan memanjatkan puja dan puji syukur kehadiran Tuhan Yang Maha Esa, yang telah memberikan rahmat yang melimpah dan kesehatan, sehingga penulis bisa menyelesaikan skripsi ini dengan judul “Analisis Sentimen Opini Pengunjung Pada Event Budaya Pop Jepang Pada Media Sosial Twitter Menggunakan Algoritma Support Vector Machine”, sebagai salah satu syarat untuk menyelesaikan program strata 1 Sarjana Ilmu Komputer Universitas AMIKOM Yogyakarta.

Selama proses menyusun skripsi ini, penulis mendapat dukungan, bantuan, bimbingan, dan nasehat dari berbagai pihak. Pada kesempatan ini penulis menyampaikan terima kasih kepada:

1. Bapak Prof. Dr. M. Suyanto, MM. Selaku Rektor Universitas AMIKOM Yogyakarta.
2. Bapak Hanif Al Fatta, S.Kom., M.Kom. Selaku Dekan Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta.
3. Bapak Yoga Pristyanto, S.Kom, M.Eng. Selaku dosen pembimbing skripsi yang telah memberikan bimbingan dan berbagi pengalaman kepada penulis.
4. Seluruh *civitas* akademika Universitas AMIKOM Yogyakarta yang telah memberikan pengetahuan dan jasanya kepada penulis selama perkuliahan.

Yogyakarta, 17 Oktober 2023

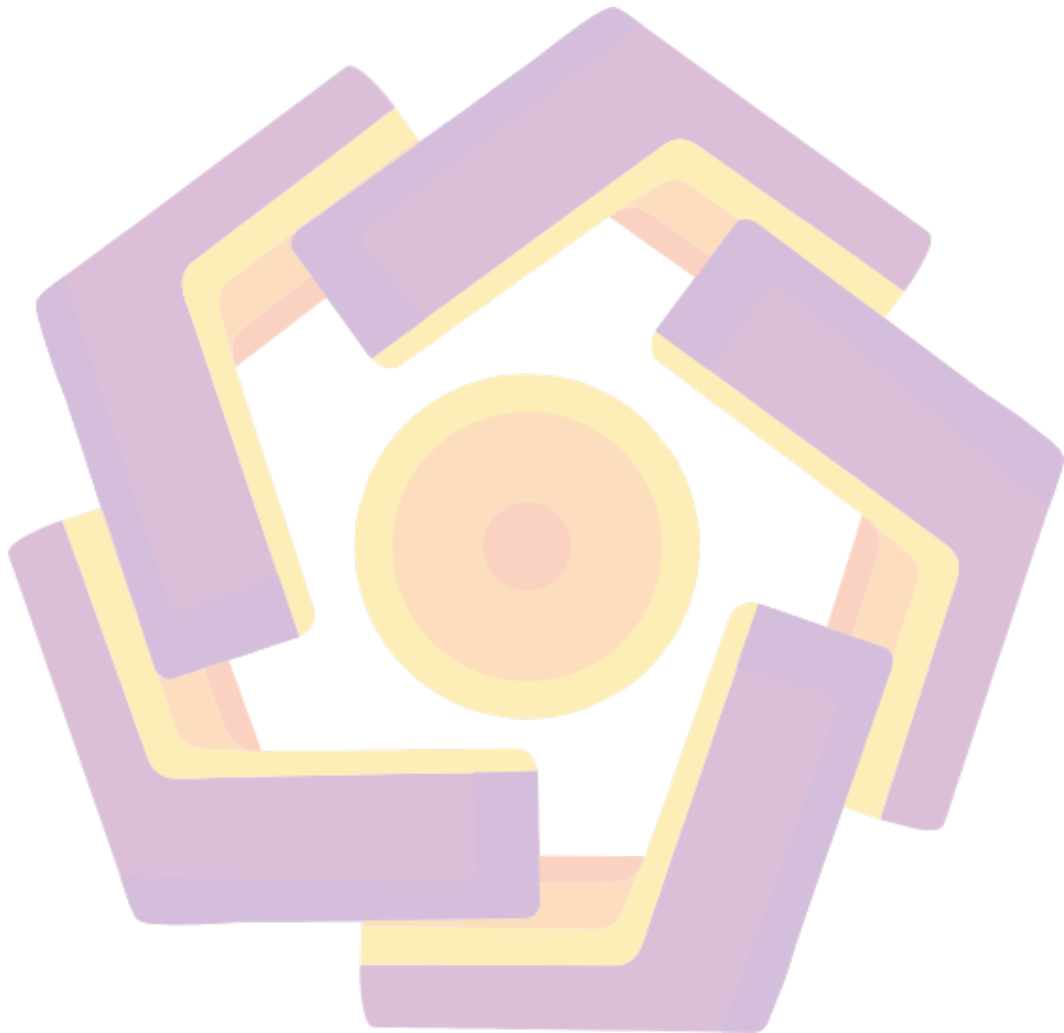
Penulis

DAFTAR ISI

SKRIPSI.....	i
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI.....	vii
DAFTAR TABEL.....	x
DAFTAR GAMBAR.....	xi
DAFTAR LAMBANG DAN SINGKATAN	xii
DAFTAR ISTILAH.....	xiii
INTISARI	xiv
ABSTRACT.....	xv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	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	4
BAB II TINJAUAN PUSTAKA	5
2.1 Studi Literatur.....	5
2.2 Dasar Teori	9
2.2.1 Support Vector Machine	9
2.2.2 Multinomial Naïve Bayes	10
2.2.3 Twitter.....	11
2.2.4 Text Mining	11
2.2.5 Analisis Sentimen	12
2.2.6 Data Labelling VADER.....	13

2.2.7	TF-IDF	13
2.2.8	Confusion Matrix	15
2.2.8	Preprocessing Data.....	16
BAB III METODE PENELITIAN		18
3.1	Alur Penelitian	18
3.1.1	Data Collection	19
3.1.2	Data Translation.....	19
3.1.3	Preprocessing Data.....	19
3.1.4	Data Labelling.....	21
3.1.5	Feature Extraction (TF-IDF).....	22
3.1.6	Splitting Data	22
3.1.7	Oversampling	23
3.1.8	Model	23
3.1.9	Evaluation	23
3.1.10	Result	25
3.2	Alat dan Bahan.....	26
BAB IV HASIL DAN PEMBAHASAN		27
4.1	Persiapan dan Data Collection.....	27
4.2	Data Translation.....	28
4.3	Preprocessing Data	30
4.4	Labelling Data.....	38
4.5	Feature Extraction (TF-IDF).....	39
4.6	Splitting Data	39
4.7	Oversampling.....	40
4.8	Classification Report Support Vector Machine	40
4.9	Classification Report Multinomial Naive Bayes	42
4.10	Hasil Akhir.....	43
BAB V PENUTUP		44

5.1 Kesimpulan	44
5.2 Saran	44
DAFTAR PUSTAKA	45

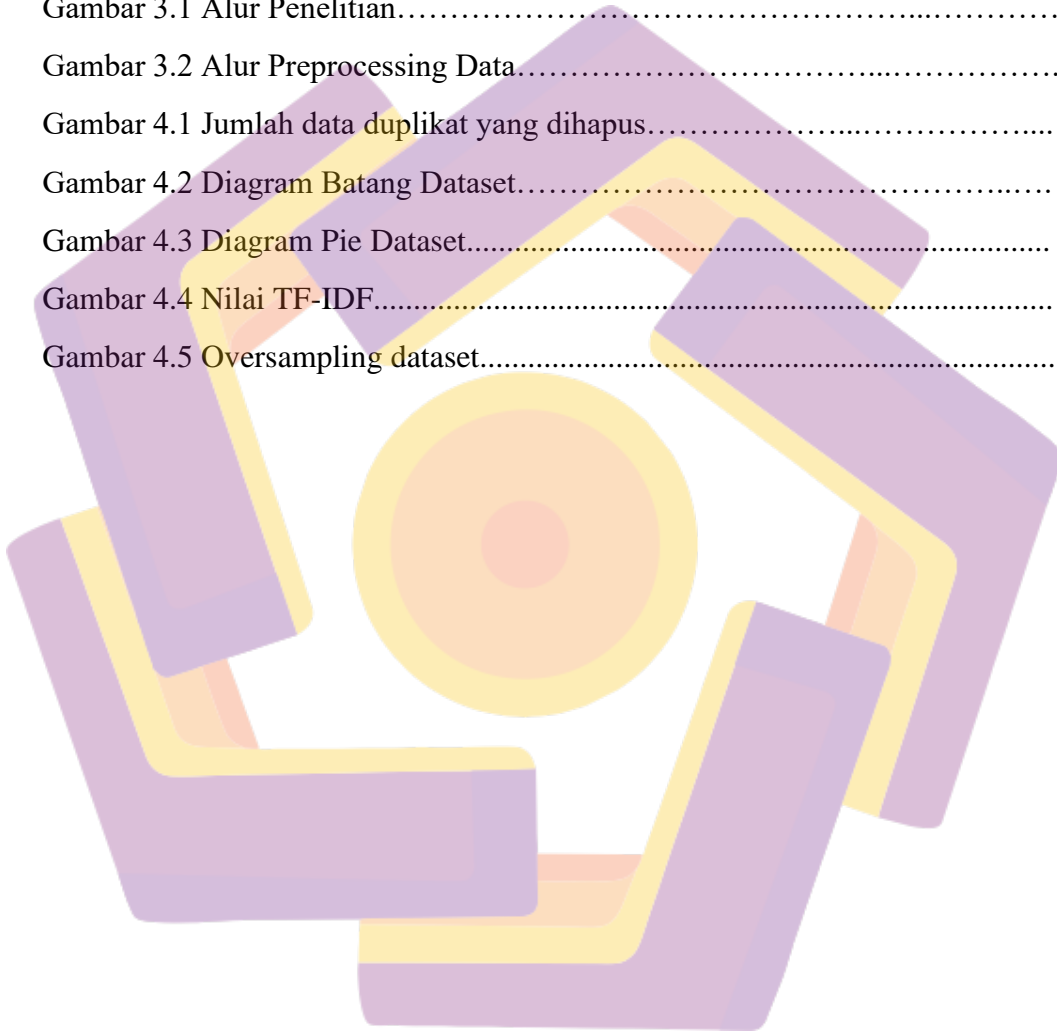


DAFTAR TABEL

Tabel 2.1 Keaslian Penelitian.....	7
Tabel 3.1 Evaluation Measure Formula.....	24
Tabel 3.2 Binary Class Confusion Matrix.....	25
Tabel 3.3 Multi-class Confusion Matrix.....	25
Tabel 4.1 Dataset hasil scraping data Twitter.....	27
Tabel 4.2 Data setelah diterjemahkan.....	29
Tabel 4.3 Data hasil drop kolom.....	30
Tabel 4.4 Data setelah cleaning.....	31
Tabel 4.5 Data duplikat.....	32
Tabel 4.6 Data hasil casefolding.....	33
Tabel 4.7 Data hasil Tokenizing.....	35
Tabel 4.8 Filtering stopwords.....	36
Tabel 4.9 Stemming menggunakan Snowball Stemmer.....	37
Tabel 4.10 Multi-class Confusion Matrix.....	41
Tabel 4.11 Confusion Matrix SVM.....	41
Tabel 4.12 Jumlah Sentimen True Prediction Menggunakan SVM.....	41
Tabel 4.13 Evaluasi Performa SVM.....	42
Tabel 4.14 Confusion Matrix MNB.....	42
Tabel 4.15 Jumlah Sentimen True Prediction Menggunakan MNB.....	42
Tabel 4.16 Evaluasi Performa MNB.....	43
Tabel 4.17 Perbandingan Metric Measure SVM dan MNB.....	43

DAFTAR GAMBAR

Gambar 2.1 Support Vector Machine.....	9
Gambar 2.2 Ekstraksi data menggunakan data scraping.....	12
Gambar 2.3 Confussion Matrix.....	15
Gambar 3.1 Alur Penelitian.....	18
Gambar 3.2 Alur Preprocessing Data.....	20
Gambar 4.1 Jumlah data duplikat yang dihapus.....	33
Gambar 4.2 Diagram Batang Dataset.....	38
Gambar 4.3 Diagram Pie Dataset.....	39
Gambar 4.4 Nilai TF-IDF.....	39
Gambar 4.5 Oversampling dataset.....	40

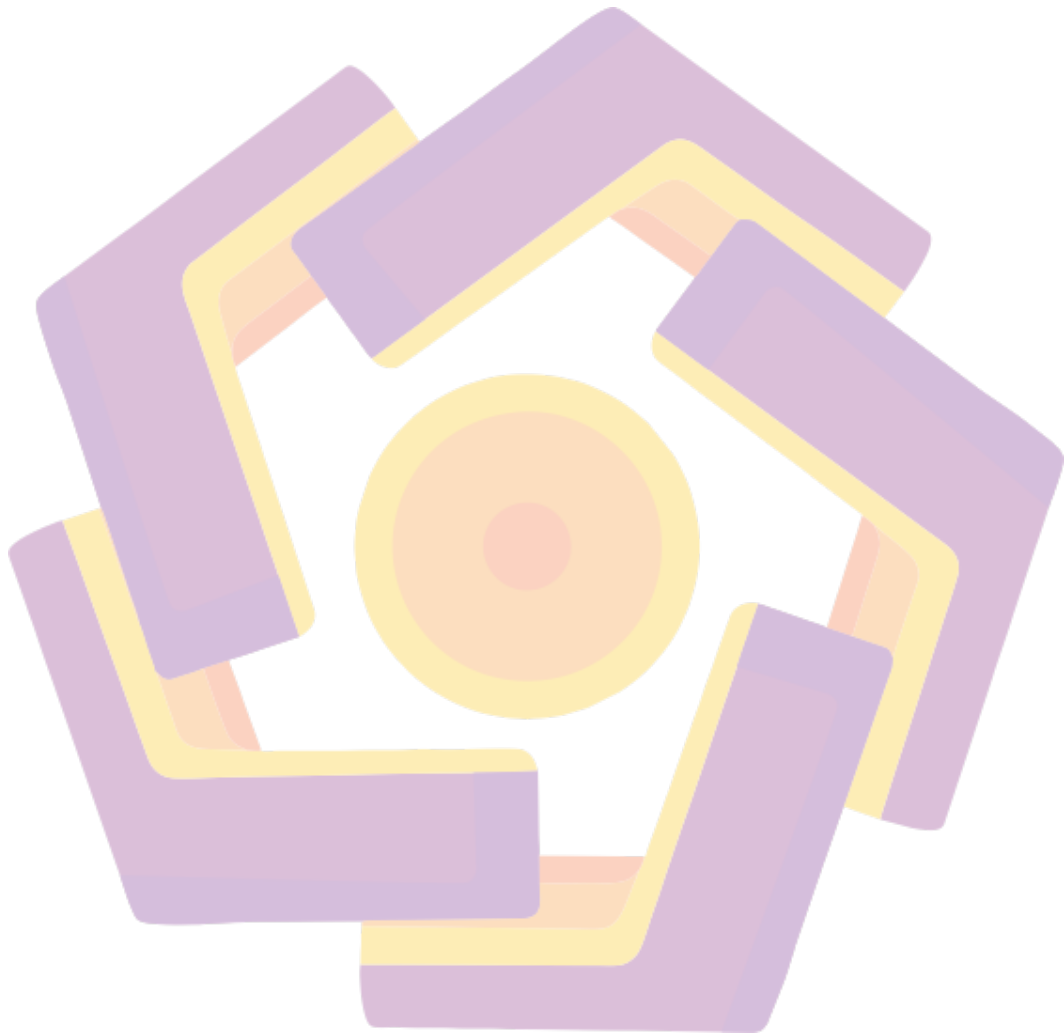


DAFTAR LAMBANG DAN SINGKATAN

SVM	Support Vector Machines
RT	Retweet
NLP	Natural Language Processing
MNB	Multinomial Naïve Bayes
VADER	Valence Aware Dictionary and Sentiment Reasoner
TF-IDF	Term Frequency-Inverse Document Frequency
TP	True Positive
TN	True Negative
FP	False Positive
FN	False Negative
TNe	True Neutral
FNe	False Neutral

DAFTAR ISTILAH

Vektor	Besaran yang mempunyai arah
Dataset	Sekumpulan data



INTISARI

Media sosial, khususnya Twitter, telah menjadi *platform* utama bagi pengunjung event budaya pop Jepang untuk berbagi pendapat dan pengalaman mereka. Event-event ini, seperti konser musik, festival anime, dan acara cosplay, memiliki penggemar yang cukup besar di Indonesia dan cukup dibicarakan di media sosial. Analisis sentimen memiliki peran penting dalam memahami respons dan sentimen dari pengunjung terhadap event-event ini. Penelitian ini menggunakan salah satu algoritma dari *Supervised Learning* yaitu *Support Vector Machine*. mampu mengklasifikasikan teks menjadi kategori-kategori sentimen seperti positif, negatif, atau netral. Dari algoritma tersebut menggunakan kernel linear dengan algoritma pembandingan yaitu *Multinomial Naïve Bayes* untuk melihat perbandingan performa pada kedua algoritma tersebut. Dari percobaan yang dilakukan dengan pembagian data latih sebesar 90% dan data latih sebesar 10% dari 3968 data, *Support Vector Machine* mendapatkan hasil akurasi sebesar 83.3% sedangkan *Multinomial Naïve Bayes* mendapatkan akurasi sebesar 68.7%. *Support Vector Machine* mendapatkan akurasi lebih baik dibandingkan *Multinomial Naïve Bayes*.

Kata kunci: Support Vector Machine, Multinomial Naïve Bayes, Analisis Sentimen, Twitter.

ABSTRACT

Social media, particularly Twitter, has become a major platform for visitors to Japanese pop culture events to share their opinions and experiences. These events, such as music concerts, anime festivals, and cosplay events, have a sizable fan base in Indonesia and are quite talked about on social media. Sentiment analysis plays an important role in understanding the responses and sentiments of visitors to these events. This research uses one of the algorithms from Supervised Learning, namely Support Vector Machine, which is able to classify text into sentiment categories such as positive, negative, or neutral. The algorithm uses a linear kernel with a comparison algorithm, namely Multinomial Naïve Bayes, to see the performance comparison of the two algorithms. From the experiments conducted with the division of training data by 90% and training data by 10% of 3968 data, Support Vector Machine gets an accuracy result of 83.3% while Multinomial Naïve Bayes gets an accuracy of 68.7%. Support Vector Machine gets better accuracy than Multinomial Naïve Bayes.

Keyword: *Support Vector Machine, Multinomial Naïve Bayes, Sentiment Analysis, Twitter.*