

**Analisis Sentimen Tweets Penggunaan Artificial Intelligence  
dalam Menghasilkan Gambar di Media Sosial X Menggunakan  
Algoritma Naive Bayes dan Support Vector Machine (SVM)**

**SKRIPSI**

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi S1 Informatika



disusun oleh

**ASNUR AMIRUDIN**

**22.21.1546**

Kepada

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

**2024**

**Analisis Sentimen Tweets Penggunaan Artificial Intelligence  
dalam Menghasilkan Gambar di Media Sosial X Menggunakan  
Algoritma Naive Bayes dan Support Vector Machine (SVM)**

**SKRIPSI**

untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi S1 Informatika



disusun oleh

**ASNUR AMIRUDIN**

**22.21.1546**

Kepada

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

**2024**

I

**HALAMAN PERSETUJUAN**

**SKRIPSI**

**Analisis Sentimen Tweets Penggunaan Artificial Intelligence dalam  
Menghasilkan Gambar di Media Sosial X Menggunakan Algoritma Naive  
Bayes dan Support Vector Machine (SVM)**

yang disusun dan diajukan oleh

**Asnur Amirudin**

**22.21.1546**

telah disetujui oleh Dosen Pembimbing Skripsi  
pada tanggal 24 Juli 2024

**Dosen Pembimbing,**

  
**Akhmad Dahlan, M.Kom**  
**NIK. 190302174**

**HALAMAN PENGESAHAN**

**SKRIPSI**

**Analisis Sentimen Tweets Penggunaan Artificial Intelligence dalam  
Menghasilkan Gambar di Media Sosial X Menggunakan Algoritma Naive  
Bayes dan Support Vector Machine (SVM)**

yang disusun dan diajukan oleh

**Asnur Amirudin**

**22.21.1546**

Telah dipertahankan di depan Dewan Penguji  
pada tanggal 24 Juli 2024

**Susunan Dewan Penguji**

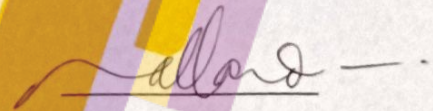
**Nama Penguji**

**Majid Rahardi, S.Kom., M.Eng**  
**NIK. 190302393**

**Rumini, M.Kom**  
**NIK. 190302246**

**Ahmad Dahlan, M.Kom**  
**NIK. 190302174**

**Tanda Tangan**



Skrripsi ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Sarjana Komputer  
Tanggal 22 Juli 2024

**DEKAN FAKULTAS ILMU KOMPUTER**



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

## HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Asnur Amirudin  
NIM : 22.21.1546

Menyatakan bahwa Skripsi dengan judul berikut:

**Analisis Sentimen Tweets Penggunaan Artificial Intelligence dalam Menghasilkan Gambar di Media Sosial X Menggunakan Algoritma Naive Bayes dan Support Vector Machine (SVM)**

Dosen Pembimbing : Akhmad Dahlan, M.Kom

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 Juli 2024

Yang Menyatakan,



Asnur Amirudin

## HALAMAN PERSEMBAHAN

Pertama-tama penulis panjatkan puji syukur kepada Tuhan Yang Maha Esa. yang telah memberikan rahmat serta hidayah sehingga dapat terselesaikannya Tugas Akhir ini. Tidak lupa juga penulis bersyukur atas hadirnya orang-orang yang berarti dan mendukung dalam penyelesaian tugas akhir ini. Oleh karena itu tugas akhir ini penulis persembahkan kepada :

1. Kedua Orang Tua, yang selalu memberikan doa, dukungan, dan kasih sayang tanpa henti.
2. Dosen-dosen S1 Informatika Universitas Amikom Yogyakarta, yang telah memberikan ilmu dan bimbingan selama masa studi.
3. Dosen Pembimbing, Bapak Akhmad Dahlan, M.Kom, yang telah memberikan arahan, masukan, dan dukungan dalam penyelesaian tugas akhir ini.
4. Rekan-rekan dan sahabat, yang telah memberikan semangat, motivasi, dan bantuan selama masa perkuliahan dan proses penulisan tugas akhir ini.
5. Semua pihak yang telah berkontribusi dalam penyelesaian tugas akhir ini, yang tidak dapat disebutkan satu per satu.

## KATA PENGANTAR

Puji Syukur penulis panjatkan kehadiran Tuhan Yang Maha Esa yang telah melimpahkan rahmat serta hidayah, sehingga penulis dapat menyelesaikan penelitian yang berjudul “Analisis Sentimen Tweets Penggunaan Artificial Intelligence dalam Menghasilkan Gambar di Media Sosial X Menggunakan Algoritma Naive Bayes dan Support Vector Machine (SVM)”. Penelitian ini ditulis sebagai salah satu syarat untuk menyelesaikan pendidikan pada Program Studi S1 Informatika Universitas Amikom Yogyakarta.

Dalam penyelesaian Penelitian ini, tidak lepas dari segala bantuan dan dukungan yang telah diberikan kepada penulis, baik berupa petunjuk, bimbingan, keterangan pendukung, serta data yang diberikan secara lisan maupun tulisan. Oleh karena itu, penulis mengucapkan banyak terima kasih kepada :

1. Bapak Prof. Dr. M. Suyanto, MM., selaku Rektor Universitas Amikom Yogyakarta.
2. Ibu Windha Mega Pradnya Duhita, M.Kom. selaku Ketua Prodi S1 Informatika.
3. Bapak Akhmad Dahlan, M.Kom. selaku dosen pembimbing.
4. Seluruh Dosen Program Studi S1 Informatika Universitas Amikom Yogyakarta.
5. Universitas AMIKOM Yogyakarta, yang telah menyediakan fasilitas dan lingkungan yang mendukung selama masa studi.

Yogyakarta, 18 Juli 2024



Asnur Amirudin

## DAFTAR ISI

<b>HALAMAN JUDUL</b>	<b>I</b>
<b>HALAMAN PERSETUJUAN</b>	<b>II</b>
<b>HALAMAN PENGESAHAN</b>	<b>III</b>
<b>HALAMAN PERNYATAAN KEASLIAN SKRIPSI</b>	<b>IV</b>
<b>HALAMAN PERSEMBAHAN</b>	<b>V</b>
<b>KATA PENGANTAR</b>	<b>VI</b>
<b>DAFTAR ISI</b>	<b>VII</b>
<b>DAFTAR TABEL</b>	<b>IX</b>
<b>DAFTAR GAMBAR</b>	<b>X</b>
<b>DAFTAR LAMBANG DAN SINGKATAN</b>	<b>XI</b>
<b>DAFTAR ISTILAH</b>	<b>XII</b>
<b>INTISARI</b>	<b>XIII</b>
<b>ABSTRACT</b>	<b>XIV</b>
<b>BAB I PENDAHULUAN</b>	<b>1</b>
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	4
<b>BAB II TINJAUAN PUSTAKA</b>	<b>5</b>
2.1 Studi Literatur	5
2.2 Dasar Teori	12
2.2.1. Text Mining	12
2.2.2. Sosial Media X	12
2.2.3. Artificial Intelligence	13
2.2.4. Analisis Sentimen	14
2.2.5. Preprocessing	14
2.2.6. Pembobotan TF-IDF	16
2.2.7. Oversampling SMOTE	17
2.2.8. Naive Bayes	17
2.2.9. Support Vector Machine	18
<b>BAB III METODE PENELITIAN</b>	<b>20</b>
3.1 Objek Penelitian	20
3.2 Alur Penelitian	20
3.2.1. Scraping Data	21



3.2.2. Preprocessing	21
3.2.2.1 Cleaning	22
3.2.2.2 Normalization	23
3.2.2.3 Stopword	23
3.2.2.4 Tokenize	24
3.2.2.5 Stemming	24
3.2.3. Labeling	25
3.2.4. TF-IDF	26
3.2.5. Splitting Data	27
3.2.6. Oversampling SMOTE	27
3.2.7. Klasifikasi Model NB dan SVM	28
3.2.8. Evaluasi	29
3.3 Alat dan Bahan	30
3.3.1. Data Penelitian	30
3.3.2. Alat/instrumen	31
<b>BAB IV HASIL DAN PEMBAHASAN</b>	32
4.1. Implementasi Pengambilan Data	32
4.2. Implementasi Preprocessing	33
4.3. Implementasi Labeling	40
4.4. Implementasi Pembobotan TF-IDF	42
4.5. Implementasi Pembagian Data	43
4.6. Implementasi Oversampling SMOTE	45
4.7. Implementasi Klasifikasi Model Naive Bayes	47
4.8. Implementasi Klasifikasi Model SVM	48
4.9. Hasil dan Evaluasi	49
<b>BAB V PENUTUP</b>	57
5.1 Kesimpulan	57
5.2 Saran	57
<b>REFERENSI</b>	59

## DAFTAR TABEL

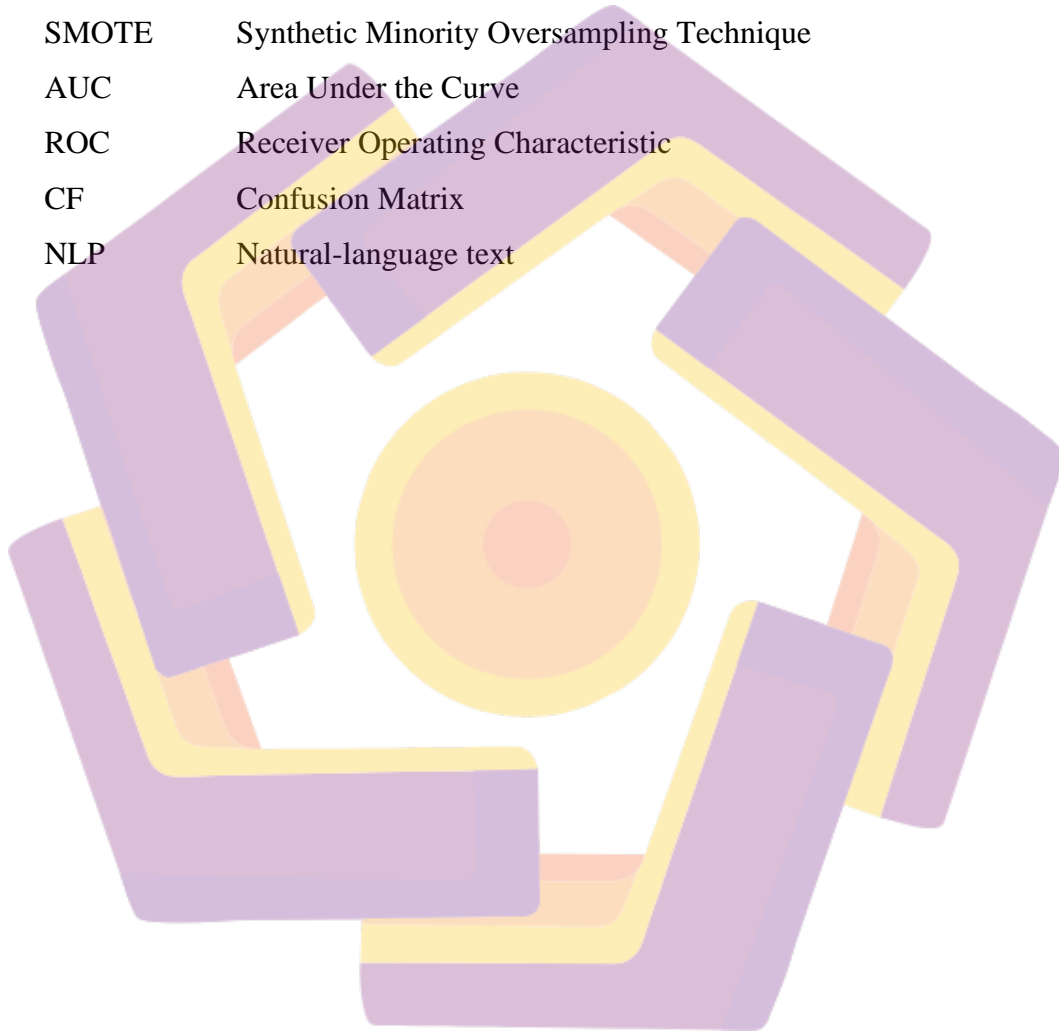
Table 2.1 Keaslian Penelitian	8
Table 3.1 Contoh Cleaning Data	22
Table 3.2 Contoh Normalization Data	23
Table 3.3 Contoh Stopword Data	24
Table 3.4 Contoh Tokenize Data	24
Table 3.5 Contoh Stemming Data	25
Table 3.6 Contoh Labeling Data	25
Table 4.1 Contoh Scraping Data	34
Table 4.2 Cleaning Data	35
Table 4.3 Normalization Data	36
Table 4.4 Stopword Data	37
Table 4.5 Tokenize Data	38
Table 4.6 Stemming Data	39
Table 4.4 Kata Teratas Setiap Kategori	41
Table 4.5 Splitting Data	44
Table 4.6 Oversampling SMOTE	46
Table 4.7 Perbandingan Accuracy	49
Table 4.8 Classification Report NB	50
Table 4.9 Classification Report SVM	51

## DAFTAR GAMBAR

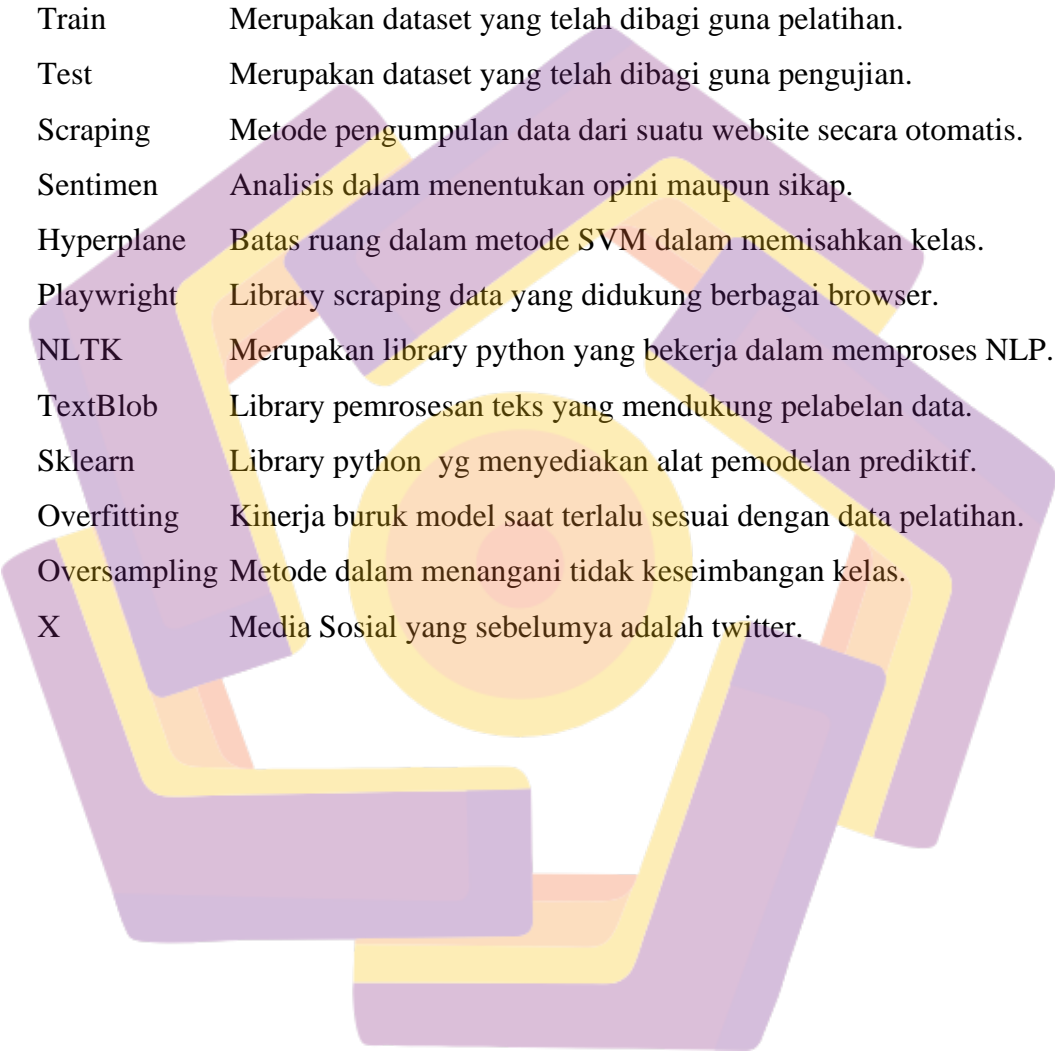
Gambar 3.1 Alur Penelitian	20
Gambar 3.2 Program Tweet-Harvest	21
Gambar 3.3 Alur Preprocessing	22
Gambar 3.4 Rumus TF	26
Gambar 3.5 Rumus IDF	26
Gambar 3.6 Rumus Score TF-IDF	27
Gambar 3.7 Rumus sampel sintetis smote	28
Gambar 3.8 Rumus Naive Bayes	28
Gambar 3.9 Rumus SVM	29
Gambar 4.1 Informasi DataFrame	33
Gambar 4.2 Labeling	40
Gambar 4.3 Visualisasi Word Cloud	42
Gambar 4.4 Code Program TF-IDF	42
Gambar 4.5 Hasil TF-IDF	43
Gambar 4.6 Code Program Pembagian Data	44
Gambar 4.7 Code Program SMOTE	45
Gambar 4.7 Code Program TF-IDF After Oversampling	46
Gambar 4.8 Code Program Naive Bayes	47
Gambar 4.9 Code Program Support Vector Machine	48
Gambar 4.10 Code Program Evaluasi Model	49
Gambar 4.11 Code Program Confusion Matrix	52
Gambar 4.12 Visualisasi Confusion Matrix NB	52
Gambar 4.13 Visualisasi Confusion Matrix SVM	53
Gambar 4.14 Visualisasi Curva ROC	54
Gambar 4.15 Perbandingan waktu pemrosesan	55
Gambar 4.16 Visualisasi Evaluasi Model	56

## DAFTAR LAMBANG DAN SINGKATAN

NB	Naive Bayes
SVM	Support Vector Machines
AI	Naive
TF-IDF	Term Frequency by Inverse Document Frequency
SMOTE	Synthetic Minority Oversampling Technique
AUC	Area Under the Curve
ROC	Receiver Operating Characteristic
CF	Confusion Matrix
NLP	Natural-language text



## DAFTAR ISTILAH



Precision	Rasio jumlah true positive dengan total jumlah prediksi positive.
Recall	Rasio jumlah true positive dengan total jumlah instance positive.
F1-Score	Merupakan rata-rata dari precision dan recall.
Train	Merupakan dataset yang telah dibagi guna pelatihan.
Test	Merupakan dataset yang telah dibagi guna pengujian.
Scraping	Metode pengumpulan data dari suatu website secara otomatis.
Sentimen	Analisis dalam menentukan opini maupun sikap.
Hyperplane	Batas ruang dalam metode SVM dalam memisahkan kelas.
Playwright	Library scraping data yang didukung berbagai browser.
NLTK	Merupakan library python yang bekerja dalam memproses NLP.
TextBlob	Library pemrosesan teks yang mendukung pelabelan data.
Sklearn	Library python yg menyediakan alat pemodelan prediktif.
Overfitting	Kinerja buruk model saat terlalu sesuai dengan data pelatihan.
Oversampling	Metode dalam menangani tidak keseimbangan kelas.
X	Media Sosial yang sebelumnya adalah twitter.

## INTISARI

Kecerdasan buatan telah berkembang sangat cepat dalam beberapa tahun terakhir hal ini telah memicu kontroversi dalam berbagai macam bidang terutama bidang seni. Maka dari itu dilakukan penelitian untuk menganalisis sentimen terhadap penggunaan kecerdasan buatan dalam menghasilkan gambar di media sosial X dengan menggunakan algoritma *Naive Bayes* dan *Support Vector Machine*. Data yang di analisis sejumlah 25.029 data yang didapatkan melalui *crawling* tweet dengan rentang Januari hingga Juni 2024. Hasil dari penelitian ini menunjukkan bahwa model *SVM* memiliki performa yang lebih baik dengan akurasi 94.40%, dibandingkan dengan model *Naive Bayes* yang memiliki akurasi 74.57%. Model *Naive Bayes* dengan *precision* sebesar 0.80, *recall* sebesar 0.75, dan *F1-Score* sebesar 0.76. Sementara itu, model *SVM* mencapai, *precision* sebesar 0.94, *recall* sebesar 0.94, dan *F1-Score* sebesar 0.94. *Area Under the Curve (AUC)* untuk *Naive Bayes* adalah 0.8589, sedangkan untuk *SVM* adalah 0.9794. *Time processing* untuk *Naive Bayes* didapatkan 0,0637 detik untuk training dan testing sebesar 0,0021 detik sedangkan *SVM* mendapatkan training 1164,1675 detik, dan testing 10,7715 detik. Hal ini menjadikan *SVM* sebagai model yang lebih efektif dalam melakukan analisis sentimen tetapi *Naive Bayes* lebih unggul untuk efisiensi waktu.

**Kata kunci:** Analisis Sentimen, Klasifikasi Sentimen, Naive Bayes, Support Vector Machine, Kecerdasan Buatan

## ***ABSTRACT***

*Artificial intelligence has developed very quickly in recent years, this has sparked controversy in various fields, especially the arts. Therefore, research was conducted to analyze sentiment regarding the use of artificial intelligence in producing images on social media X using the Naive Bayes algorithm and Support Vector Machine. The data analyzed was 25,029 data obtained through crawling tweets ranging from January to June 2024. The results of this research show that the SVM model has better performance with an accuracy of 94.40%, compared to the Naive Bayes model which has an accuracy of 74.57%. Naive Bayes model with precision of 0.80, recall of 0.75, and F1-Score of 0.76. Meanwhile, the SVM model achieved precision of 0.94, recall of 0.94, and F1-Score of 0.94. The Area Under the Curve (AUC) for Naive Bayes is 0.8589, while for SVM it is 0.9794. The processing time for Naive Bayes was 0.0637 seconds for training and testing of 0.0021 seconds, while SVM got 1164.1675 seconds for training and 10.7715 seconds for testing. This makes SVM a more effective model in conducting sentiment analysis but Naive Bayes is superior for time efficiency.*

***Keyword:*** *Sentiment Analysis, Sentiment Classification, Naive Bayes, Support Vector Machine, Artificial Intelligence*