

**KLASIFIKASI ANALISIS SENTIMEN PADA ULASAN  
PENGGUNA APLIKASI “THREADS” DI GOOGLE PLAY  
STORE MENGGUNAKAN ALGORITMA NAÏVE BAYES DAN  
SUPPORT VECTOR MACHINE DENGAN SMOTE**

**SKRIPSI**

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



disusun oleh

**SENDI TIANTARA**

**20.11.3473**

Kepada

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

**2024**

**KLASIFIKASI ANALISIS SENTIMEN PADA ULASAN  
PENGGUNA APLIKASI “THREADS” DI GOOGLE PLAY  
STORE MENGGUNAKAN ALGORITMA NAÏVE BAYES DAN  
SUPPORT VECTOR MACHINE DENGAN SMOTE**

**SKRIPSI**

untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi (*SI informatika*)



disusun oleh

**SENDI TIANTARA**

**20.11.3473**

Kepada

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

**2024**

**HALAMAN PERSETUJUAN**

**SKRIPSI**

**KLASIFIKASI ANALISIS SENTIMEN PADA ULASAN  
PENGGUNA APLIKASI “THREADS” DI GOOGLE PLAY  
STORE MENGGUNAKAN ALGORITMA NAÏVE BAYES DAN  
SUPPORT VECTOR MACHINE DENGAN SMOTE**

yang disusun dan diajukan oleh

**Sendi tiantara**

**20.11.3473**

telah disetujui oleh Dosen Pembimbing Skripsi  
pada tanggal 22 Agustus 2024

Dosen Pembimbing,

  
**Mulia Sulistiyono, M.Kom**

**NIK. 190302248**

**HALAMAN PENGESAHAN  
SKRIPSI**

**KLASIFIKASI ANALISIS SENTIMEN PADA ULASAN  
PENGGUNA APLIKASI “THREADS” DI GOOGLE PLAY  
STORE MENGGUNAKAN ALGORITMA NAÏVE BAYES DAN  
SUPPORT VECTOR MACHINE DENGAN SMOTE**

yang disusun dan diajukan oleh

**Sendi tiantara**

**20.11.3473**

Telah dipertahankan di depan Dewan Penguji  
pada tanggal 22 Agustus 2024

**Susunan Dewan Penguji**

**Nama Penguji**

**Tanda Tangan**

**Alfie Nur Rahmi, M.Kom**  
**NIK. 190302240**

**Arifiyanto Hadinegoro, S.Kom., M.T**  
**NIK . 190302289**

**Mulia Sulistiyono, M.Kom**  
**NIK. 190302248**



Skripsi ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Sarjana Komputer  
Tanggal 22 Agustus 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 : Sendi tiantara**  
**NIM : 20.11.3473**

Menyatakan bahwa Skripsi dengan judul berikut:

**KLASIFIKASI ANALISIS SENTIMEN PADA ULASAN  
PENGGUNA APLIKASI “THREADS” DI GOOGLE PLAY  
STORE MENGGUNAKAN ALGORITMA NAÏVE BAYES DAN  
SUPPORT VECTOR MACHINE DENGAN SMOTE**

Dosen Pembimbing : Mulia Sulistiyono, 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, 22 Agustus 2024

Yang Menyatakan,



Sendi tiantara



## HALAMAN PERSEMBAHAN

Alhamdulillah segala puji dan syukur bagi Allah SWT dan segala rahmat serta hidayahnya sehingga penulis dapat menyelesaikan skripsi ini dengan baik, tidak terlepas dari beberapa pihak yang tulus membantu secara langsung maupun tidak langsung sehingga penulis dapat menyelesaikan skripsi ini dengan baik. Oleh karena itu saya ucapkan terimakasih dan skripsi ini saya persembahkan kepada :

1. Ayah penulis yang selalu menjadi Panutan, Bapak Bastiar. Beliau yang selalu memotivasi, memberikan dukungan hingga penulis mampu menyelesaikan studi sampai mendapatkan gelar sarjana. Beliau memang tidak sempat merasakan Pendidikan sampai bangku perkuliahan, namun beliau mampu menjadikan anak semata wayangnya ini merasakan tahap Pendidikan hingga ke jenjang perguruan tinggi. Terimakasih atas segala dan kasih sayang, pengorbanan, dan doanya yang tiada henti mengalir untuk keberhasilan penulis.
2. Pintu surgaku, Ibu Rosmawati, beliau memang juga memang tidak sempat merasakan Pendidikan sampai bangku perkuliahan, namun semangat, motivasinya sangat berperan penting dalam setiap Langkah yang penulis ambil. Dukungan, kasih sayang, dan pengorbanan tidak ternilai harganya. Ibu selalu memberikan semangat dan doa yang tiada henti, menjadi sumber kekuatan penulis dalam menyelesaikan studi ini. Terima kasih atas cinta dan ketulusan yang Ibu berikan.
3. Kepada Bapak Mulia Sulistiyono, S.Kom M.Kom, Terimakasih karena telah sabar dan iklas dalam membimbing penulis dalam penulisan skripsi ini.
4. Serta ucapan terimakasih banyak kepada teman-teman 20-IF-03, atas semua kebersamaannya, kerjasamanya, dan dukungan yang selalu ada. Semoga tali silaturahmi kita yang terjalin selama kurang lebih 4 tahun ini tetap terus kokoh.

5. Kepada sahabat-sahabat yang senantiasa memberi dukungan, kebersamaan, dan kerjasamanya selama ini terkhusus kepada Adhe Putra Kusuma Dwi Wardhana, Mario Sebastian Carlosiano, Paskalis Nani Randang dan masih banyak lagi yang tidak bisa saya sebutkan satu persatu.
6. Tidak lupa juga untuk teman-teman kost GND yang selalu memberikan dukungan, kebersamaan, dan keceriaan selama masa studi. Terima kasih atas bantuan, motivasi, dan semangat yang kalian berikan. Kebersamaan kita akan selalu menjadi kenangan berharga dalam hidup penulis.



## KATA PENGANTAR

Puji dan sukur penulis panjatkan kepada Tuhan yang Maha Esa yang telah memberikan Rahmat, hidayah dan kekuatan sehingga saya dapat menyelesaikan skripsi yang berjudul **KLASIFIKASI ANALISIS SENTIMEN PADA ULASAN PENGGUNA APLIKASI “THREADS” DI GOOGLE PLAY STORE MENGGUNAKAN ALGORITMA NAÏVE BAYES DAN SUPPORT VECTOR MACHINE DENGAN SMOTE.**

Dalam penyusunan skripsi ini penulis banyak mengalami hambatan, namun berkat bantuan dan bimbingan dari berbagai pihak akhirnya skripsi ini dapat terselesaikan dengan baik. Penulis menyadari tanpa bantuan dan dukungan dalam penyusunan skripsi ini, skripsi ini tidak akan tersusun seperti sekarang ini. Sehingga dengan penuh kerendahan hati dan rasa hormat saya mengucapkan terimakasih kepada :

1. Kepada bapak Prof. Dr. M. Suyanto, M.M selaku Rektor Universitas Amikom Yogyakarta.
2. Kepada bapak Hanif Al Fatta, S.Kom., M.Kom selaku Dekan Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.
3. Kepada ibu Windha Mega Pradnya Duhita, M.Kom selaku Ketua Program Studi Informatika.
4. Kepada bapak Mulia Sulistiyono, M.Kom selaku dosen pembimbing.
5. Kepada bapak dan ibu dosen Universitas Amikom Yogyakarta yang telah mengajarkan ilmu pengetahuan selama di bangku perkuliahan.

Dalam penulisan skripsi ini masih banyak kekurangan yang harus di sempurnakan, sehingga penulis mengharapkan kritik dan saran pembaca terhadap skripsi ini, karena tidak mungkin skripsi ini tersusun dengan kondisi yang sempurna. Tulisan ini semoga bisa bermanfaat dan mendorong kita untuk melakukan penelitian yang lebih baik.

Yogyakarta, 22 Agustus 2024

Penulis

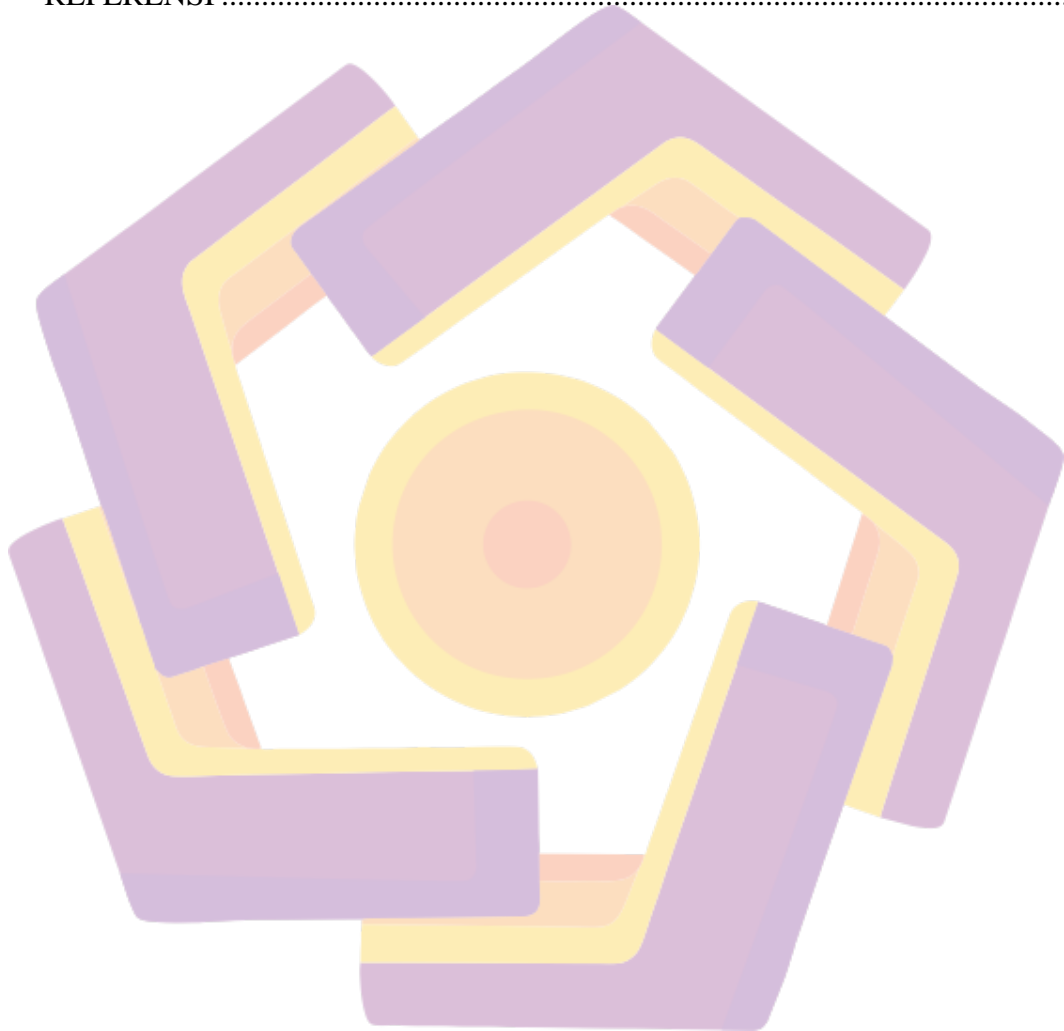


## DAFTAR ISI

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 LAMBANG DAN SINGKATAN .....	xiii
DAFTAR ISTILAH .....	xiv
INTISARI .....	xv
ABSTRACT.....	xvi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	3
1.3 Batasan Masalah.....	3
1.4 Tujuan Penelitian.....	3
1.5 Manfaat Penelitian.....	4
1.6 Sistematika Penulisan.....	4
BAB II TINJAUAN PUSTAKA .....	6
2.1 Studi Literatur .....	6

2.2	Dasar Teori .....	15
2.2.1.	Analisis sentimen .....	15
2.2.2.	Naïve Bayes .....	15
2.2.3.	Support Vector Machine .....	15
2.2.4.	Synthetic Minority Oversampling Technique (SMOTE) .....	17
2.2.5.	Confusion Matrix .....	17
2.2.6.	<i>Google colaboratory</i> (Colab).....	19
<b>BAB III METODE PENELITIAN .....</b>		<b>20</b>
3.1	Objek Penelitian .....	20
3.2	Alur Penelitian.....	20
3.2.1	Pengumpulan Data (Dari Jurnal) .....	21
3.2.2	Analisis Data .....	21
3.2.3	Perancangan Solusi .....	21
3.2.4	Analisis Hasil .....	31
3.2.5	Dokumentasi .....	31
3.3	Alat dan Bahan .....	31
<b>BAB IV HASIL DAN PEMBAHASAN .....</b>		<b>33</b>
4.1	Pengumpulan data .....	33
4.2	Preprocessing .....	33
4.3	Pelebelan data ( <i>Labelling</i> ).....	40
4.4	SMOTE .....	41
4.5	Implementasi Algoritma.....	43
4.5.1	Implementasi Naïve Bayes.....	43
4.5.2	Implementasi Support vector machine (SVM) .....	48
4.5.3	Perbandingan hasil algoritma tanpa dan dengan SMOTE .....	52

4.6	Visualisasi Word Cloud .....	53
BAB V PENUTUP .....		60
5.1	Kesimpulan.....	60
5.2	Saran.....	61
REFERENSI .....		63

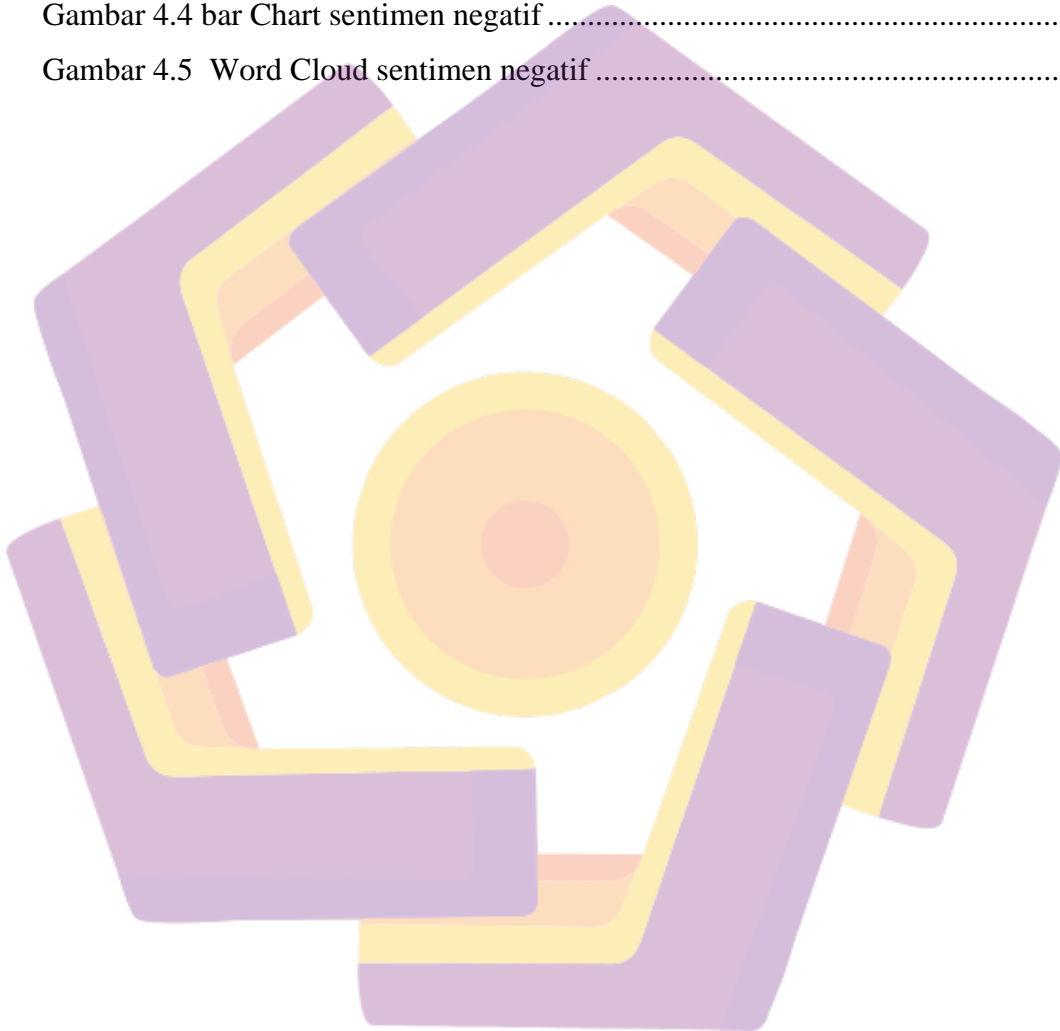


## DAFTAR TABEL

Tabel 2.1 Keaslian Penelitian .....	10
Tabel 2.2 Confusion Matrix .....	17
Tabel 3.1 Pengumpulan Data .....	22
Tabel 4.1 Source code Cleaning .....	34
Tabel 4.2 Hasil Data Cleaning .....	34
Tabel 4.3 Source code Case Folding.....	34
Tabel 4.4 Hasil Case Folding.....	35
Tabel 4.5 Source code Tokenizing.....	36
Tabel 4.6 Hasil Tokenizing.....	36
Tabel 4.7 Source code Stopword Removal.....	37
Tabel 4.8 Hasil Stopwords Removal .....	37
Tabel 4.9 Source code Stemming .....	38
Tabel 4.10 Hasil Stemming.....	38
Tabel 4.12 Hasil Pembobotan TF-IDF.....	39
Tabel 4.13 Source code labelling .....	40
Tabel 4.14 Hasil Labelling.....	40
Tabel 4.15 Nilai sebelum SMOTE.....	42
Tabel 4.16 Source code Naïve Bayes.....	43
Tabel 4.17 Hasil Confusion Matrix data uji Naïve Bayes .....	44
Tabel 4.18 Source code Naïve Bayes + SMOTE.....	45
Tabel 4.19 Hasil Confusion Matrix data uji Naïve Bayes+Smote.....	45
Tabel 4.20 Hasil perbandingan model Naïve Bayes dan Naïve Bayes + SMOTE .....	47
Tabel 4.21 Source code Support vector machine.....	48
Tabel 4.22 Hasil Confusion Matrix data uji SVM.....	48
Tabel 4.23 Source code Support vector machine + SMOTE.....	50
Tabel 4.24 Hasil Confusion Matrix data uji SVM+SMOTE .....	50
Tabel 4.25 Hasil perbandingan model Support vector machine dan Support vector machine + SMOTE .....	51
Tabel 4.27 Source code visualisasi Positif dan Negatif.....	54

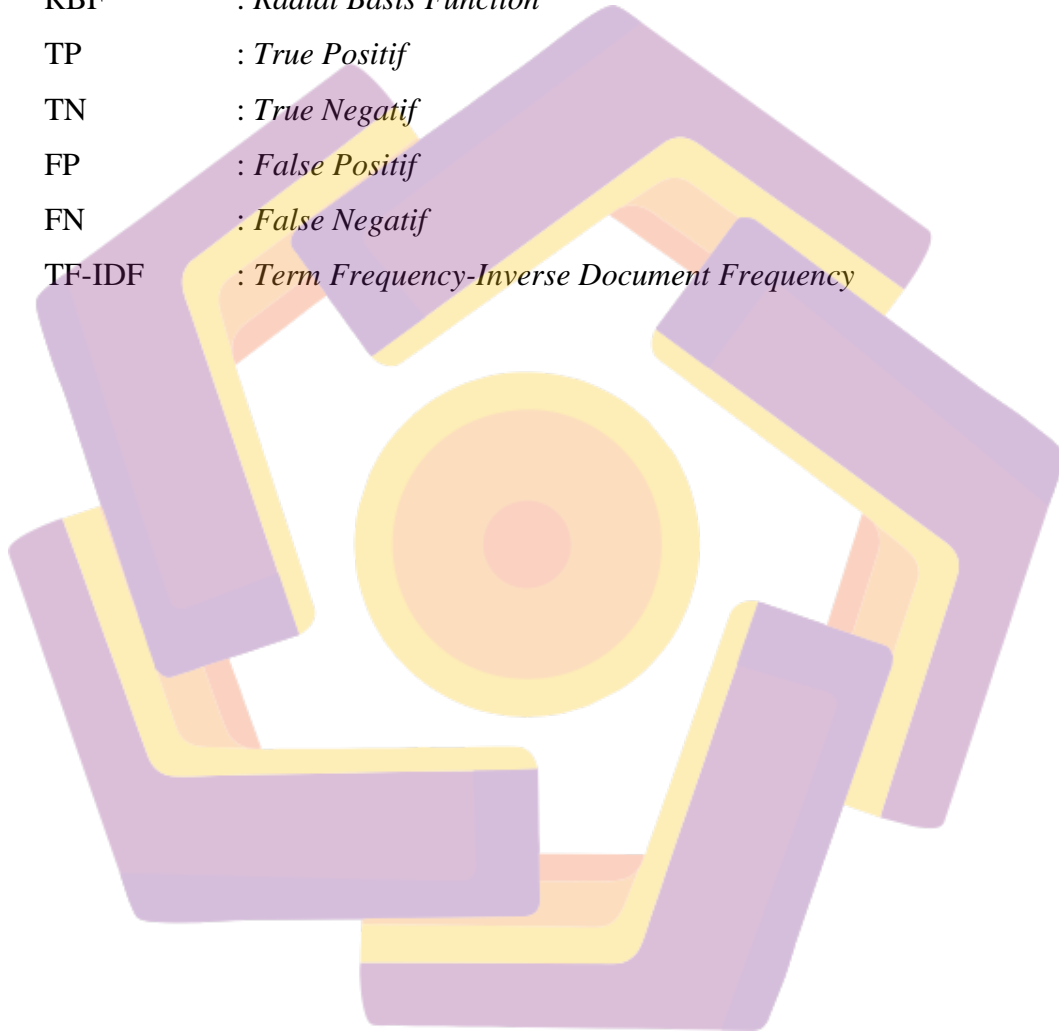
## DAFTAR GAMBAR

Gambar 2.2 Konsep dasar SVM dalam menentukan hyperplane .....	16
Gambar 3.1. Alur Penelitian .....	20
Gambar 4.2 bar Chart sentimen positif .....	55
Gambar 4.3 Word Cloud sentiment positif .....	56
Gambar 4.4 bar Chart sentimen negatif .....	57
Gambar 4.5 Word Cloud sentimen negatif .....	58



## DAFTAR LAMBANG DAN SINGKATAN

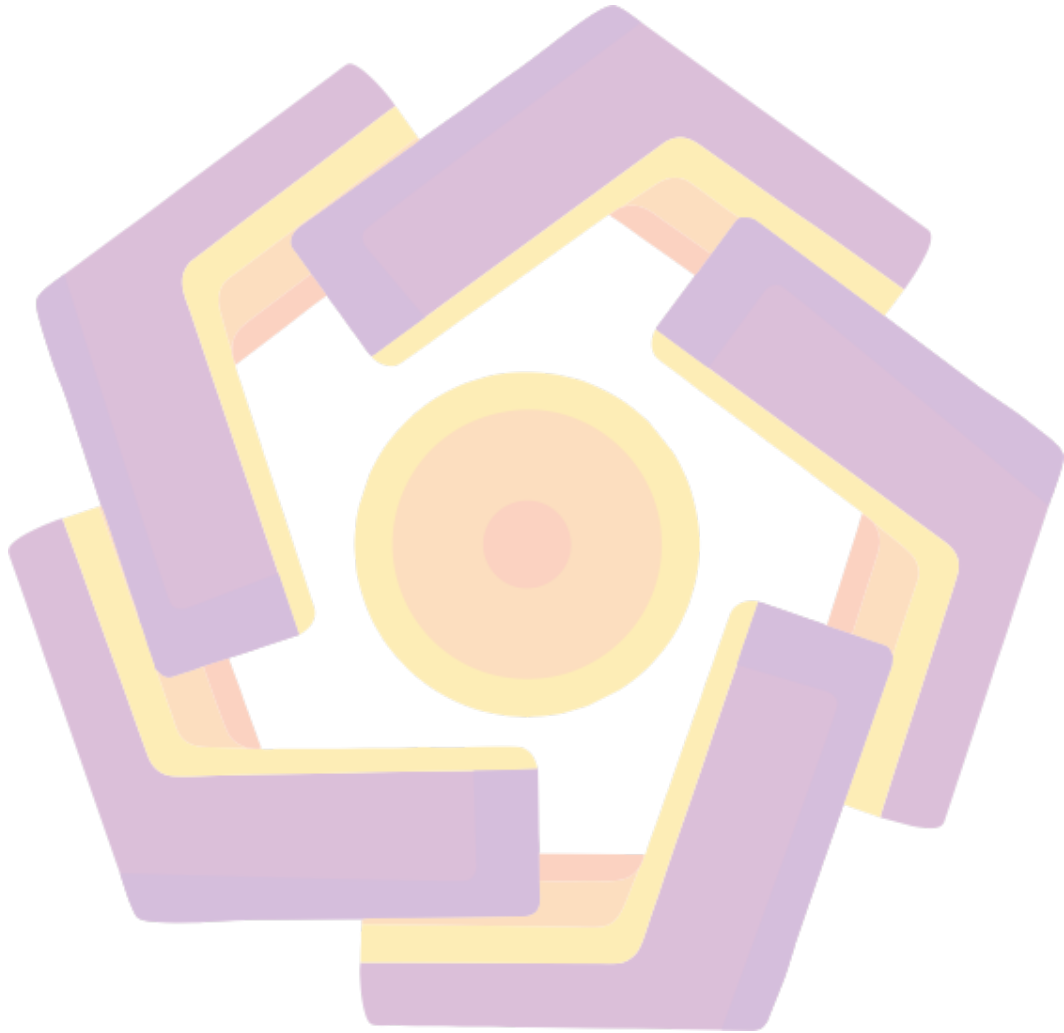
NB	: <i>Naïve bayes</i>
SVM	: <i>Support Vector Machine</i>
SMOTE	: <i>Synthetic Minority Oversampling Technique</i>
RBF	: <i>Radial Basis Function</i>
TP	: <i>True Positif</i>
TN	: <i>True Negatif</i>
FP	: <i>False Positif</i>
FN	: <i>False Negatif</i>
TF-IDF	: <i>Term Frequency-Inverse Document Frequency</i>





## DAFTAR ISTILAH

- Dataset : Sekumpulan data  
Analisis sentimen : pengenalan dan ekstraksi opini dari teks



## INTISARI

Threads, sebuah aplikasi jejaring sosial yang mendapatkan sorotan pada hari peluncurannya pada tanggal 5 juni 2023 oleh META yang dikembangkan untuk pengguna instagram. karena sudah diunduh lebih dari 100 juta kali dengan ulasan sebanyak 320 ribu pada Google Play Store. Penelitian ini bertujuan untuk mengklasifikasikan analisis sentimen pada ulasan pengguna aplikasi Threads di Google Play Store menggunakan algoritma *Naïve Bayes* dan *Support Vector Machine* (SVM) dengan teknik oversampling SMOTE. Data ulasan dikumpulkan menggunakan *Octoparse* dan terdiri dari 6.000 ulasan, dengan 3.765 ulasan dikategorikan sebagai *sentimen positif* dan 2.235 sebagai *sentimen negatif*. Ulasan dengan rating 5 dan 4 diklasifikasikan sebagai positif, sementara rating 1 dan 2 sebagai negatif. Ulasan dengan rating 3, yang sering kali bersifat netral atau campuran, dimasukkan ke dalam kategori yang paling sesuai berdasarkan konteks ulasan, yaitu negatif. Tahap *preprocessing* mencakup *cleaning*, *case folding*, *stopword removal*, dan *TF-IDF*.

Hasil penelitian menunjukkan bahwa penerapan algoritma *Naïve Bayes* dan *SVM* memberikan hasil yang berbeda dalam klasifikasi ulasan. *Naïve Bayes* menunjukkan akurasi sebesar 79.3%, precision 80.2%, recall 59.5%, dan F1-score 34.1%. Setelah penerapan SMOTE, akurasi menurun menjadi 78.0%, precision 77.9%, recall meningkat menjadi 79.1%, dan F1-score meningkat signifikan menjadi 78.4%. Sementara itu, SVM dengan kernel RBF menunjukkan kinerja yang lebih baik dengan akurasi 78.9%, precision 77.4%, recall 61.7%, dan F1-score 68.6%. Setelah penerapan SMOTE, akurasi meningkat menjadi 81.3%, precision 83.9%, recall 78.2%, dan F1-score 80.9%. hasil ini memberikan Gambaran komprehensif terkait respon pengguna terhadap aplikasi Threads, dapat menjadi panduan bagi pengembang untuk meningkatkan kualitas aplikasi

**Kata kunci:** Analisis Sentimen, *Naïve Bayes*, *Support Vector Machine*, SMOTE, Threads.

## ABSTRACT

Threads, a social networking application that gained attention on its launch day on 5 June 2023 by META developed for Instagram users. because it has been downloaded more than 100 million times with 320 thousand reviews on the Google Play Store. This research aims to classify sentiment analysis on user reviews of the Threads application on the Google Play Store using the Naïve Bayes algorithm and Support Vector Machine (SVM) with SMOTE oversampling technique. The review data was collected using Octoparse and consisted of 6,000 reviews, with 3,765 reviews categorised as positive sentiment and 2,235 as negative sentiment. Reviews with ratings of 5 and 4 were classified as positive, while ratings of 1 and 2 as negative. Reviews with a rating of 3, which are often neutral or mixed, were put into the most appropriate category based on the context of the review, which is negative. The preprocessing stage includes cleaning, case folding, stopword removal, and TF-IDF.

The results showed that the application of Naïve Bayes and SVM algorithms gave different results in review classification. Naïve Bayes showed 79.3% accuracy, 80.2% precision, 59.5% recall, and 34.1% F1-score. After applying SMOTE, the accuracy decreased to 78.0%, precision 77.9%, recall increased to 79.1%, and F1-score increased significantly to 78.4%. Meanwhile, SVM with RBF kernel showed better performance with 78.9% accuracy, 77.4% precision, 61.7% recall, and 68.6% F1-score. After the application of SMOTE, the accuracy increased to 81.3%, precision 83.9%, recall 78.2%, and F1-score 80.9%. These results provide a comprehensive overview of user responses to the Threads application, which can serve as a guide for developers to improve the quality of the application.

**Keyword:** Sentiment Analysis, Naïve Bayes, Support Vector Machine, SMOTE, Threads.