

**PERBANDINGAN ALGORITMA *SUPPORT VECTOR MACHINE*  
DAN *NAÏVE BAYES CLASSIFIER* PADA ANALISIS SENTIMEN  
APLIKASI TIKTOK MENGGUNAKAN DATA TWITTER**

**SKRIPSI**

*Diajukan sebagai salah satu persyaratan guna memperoleh gelar  
Sarjana Statistika*



**Oleh  
DENIA PUTRI FAJRINA  
NIM 19337025**

**PROGRAM STUDI SARJANA STATISTIKA  
DEPARTEMEN STATISTIKA  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS NEGERI PADANG  
2023**

## SURAT PERNYATAAN TIDAK PLAGIAT

Saya yang bertanda tangan di bawah ini:

Nama : Denia Putri Fajrina  
NIM/TM : 19337025/2019  
Program Studi : S1 Statistika  
Departemen : Statistika  
Fakultas : Matematika dan Ilmu Pengetahuan Alam

Dengan ini menyatakan, bahwa Skripsi saya dengan judul "**Perbandingan Algoritma Support Vector Machine dan Naïve Bayes Classifier Pada Analisis Sentimen Aplikasi Tiktok Menggunakan Data Twitter**" adalah benar merupakan hasil karya saya dan bukan merupakan plagiat dari karya orang lain atau pengutipan dengan cara-cara yang tidak sesuai dengan etika yang berlaku dalam tradisi keilmuan. Apabila suatu saat terbukti saya melakukan plagiat maka saya bersedia diproses dan menerima sanksi akademis maupun hukum sesuai dengan hukum dan ketentuan yang berlaku, baik di institusi UNP maupun di masyarakat dan Negara.

Demikianlah pernyataan ini saya buat dengan penuh kesadaran rasa tanggung jawab sebagai anggota masyarakat ilmiah.

Diketahui oleh,  
Kepala Departemen Statistika,



Dr. Yenni Kurniawati, S.Si., M.Si  
NIP 198402232010122005

Saya yang menyatakan,



Denia Putri Fajrina  
NIM 19337025

## PERSETUJUAN SKRIPSI

### PERBANDINGAN ALGORITMA *SUPPORT VECTOR MACHINE* DAN *NAÏVE BAYES CLASSIFIER* PADA ANALISIS SENTIMEN APLIKASI TIKTOK MENGGUNAKAN DATA TWITTER

Nama : Denia Putri Fajrina  
NIM : 19337025  
Program Studi : S1 Statistika  
Departemen : Statistika  
Fakultas : Matematika dan Ilmu Pengetahuan Alam

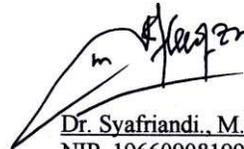
Padang, 06 November 2023

Mengetahui:  
Ketua Departemen Statistika



Dr. Yenni Kurniawati., S.Si, M.Si  
NIP. 19840223 201012 2 005

Disetujui Oleh:  
Pembimbing



Dr. Syafriandi., M.Si  
NIP. 196609081991031003

## PENGESAHAN LULUS UJIAN SKRIPSI

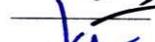
Nama : Denia Putri Fajrina  
NIM : 19337025  
Program Studi : S1 Statistika  
Departemen : Statistika  
Fakultas : Matematika dan Ilmu Pengetahuan Alam

### **PERBANDINGAN ALGORITMA *SUPPORT VECTOR MACHINE* DAN *NAÏVE BAYES CLASSIFIER* PADA ANALISIS SENTIMEN APLIKASI TIKTOK MENGGUNAKAN DATA TWITTER**

Dinyatakan lulus setelah dipertahankan di depan Tim Penguji Skripsi  
Departemen Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam  
Universitas Negeri Padang

Padang, 06 November 2023

Tim Penguji

	Nama	Tanda Tangan
Ketua	: Dr. Syafriandi, M.Si	
Anggota	: Dra. Nonong Amalita, M.Si	
Anggota	: Admi Salma, S.Pd, M.Si	

# Perbandingan Algoritma *Support Vector Machine* dan *Naïve Bayes Classifier* Pada Analisis Sentimen Aplikasi Tiktok Menggunakan Data Twitter

Denia Putri Fajrina

## ABSTRAK

Sosial media yang cukup populer saat ini adalah TikTok. Indonesia menempati peringkat kedua dengan pengguna aktif terbanyak pada awal tahun 2023. Dalam setiap produk yang memiliki kepopuleritasan sangat banyak opini masyarakat terhadap aplikasi TikTok. Tidak semua masyarakat khususnya di Indonesia dapat menerima aplikasi TikTok dengan baik, dan terdapat pro kontra terhadap aplikasi TikTok di media sosial terutama Twitter.

Analisis sentimen mengenai persepsi masyarakat terhadap aplikasi TikTok sangat penting bagi pengembang aplikasi TikTok. Melalui sentimen positif dan negatif tersebut, pengembang aplikasi dapat melakukan evaluasi untuk meningkatkan kualitas dari TikTok agar menarik bagi pengguna lama maupun pengguna baru. Untuk menganalisis sentimen dari ulasan mengenai aplikasi TikTok, digunakan metode klasifikasi algoritma *Support Vector Machine* (SVM) dan *Naïve Bayes Classifier* (NBC). Pada algoritma SVM mengandalkan sebuah *hyperplane* dengan menambahkan fungsi kernel sedangkan NBC menggunakan probabilitas bersyarat dalam melakukan klasifikasi. Penelitian ini bertujuan untuk membandingkan performa Algoritma *Support Vector Machine* (SVM) dan *Naïve Bayes Classifier* (NBC) pada analisis sentimen Aplikasi TikTok menggunakan data Twitter. Hasil yang diharapkan dari penelitian ini yaitu pertama mengetahui perbandingan akurasi dari algoritma SVM dan NBC. Kedua mengetahui penciri pada setiap sentimen positif dan negatif.

Hasil performa model dari algoritma SVM dan NBC menggunakan *confusion matrix* pada penelitian ini menunjukkan bahwa metode SVM lebih baik dalam melakukan klasifikasi sentimen aplikasi TikTok dibandingkan dengan NBC. Hal ini dapat dilihat dari hasil yang diperoleh dengan nilai akurasi SVM adalah 77,887% dan nilai akurasi NBC adalah 72,115%. Klasifikasi dan visualisasi didapatkan pengetahuan berupa kata yang sering muncul di ulasan positif seperti seperti “tiktok”, “video”, “iya”, “suka”, ”bagus”, “ nonton” dan ulasan negatif terdapat kata seperti “aplikasi”, “ konten”, “salah”, “aneh” , “gila”.

Kata Kunci : Analisis Sentimen, NBC, SVM, TikTok

# **Comparison of Support Vector Machine and Naïve Bayes Classifier Algorithms on Sentiment Analysis of Tiktok Application Using Twitter Data**

**Denia Putri Fajrina**

## **ABSTRACT**

The most popular social media today is TikTok. Indonesia ranks second with the most active users in early 2023. In every product that has popularity, there are many public opinions on the TikTok application. Not all people, especially in Indonesia, can accept the TikTok application well, and there are pros and cons to the TikTok application on social media, especially Twitter.

Sentiment analysis regarding public perception of the TikTok application is very important for TikTok application developers. Through these positive and negative sentiments, application developers can evaluate to improve the quality of TikTok to appeal to old and new users. To analyze the sentiment of reviews about the TikTok application, the classification methods of the Support Vector Machine (SVM) and Naïve Bayes Classifier (NBC) algorithms are used. The SVM algorithm relies on a hyperplane by adding a kernel function while NBC uses conditional probability in classification. This study aims to compare the performance of the Support Vector Machine (SVM) and Naïve Bayes Classifier (NBC) algorithms on sentiment analysis of the TikTok application using Twitter data. The expected results of this study are first to know the accuracy comparison of the SVM and NBC algorithms. Second, knowing the characteristics of each positive and negative sentiment.

The results of the model performance of the SVM and NBC algorithms using the confusion matrix in this study show that the SVM method is better at classifying the sentiment of the TikTok application compared to NBC. This can be seen from the results obtained with the SVM accuracy value is 77,887%. and the NBC accuracy value is 72,115%. The classification and visualization, knowledge is obtained in the form of words that often appear in positive reviews such as "tiktok", "video", "yes", "like", "good", "watch", and negative reviews there are words such as "application", "content", "wrong", "content", "crazy".

Keywords: NBC, Sentiment Analysis, SVM, TikTok

## KATA PENGANTAR

Puji syukur penulis ucapkan kepada Allah SWT, atas rahmat dan karunianya yang telah dilimpahkannya. Penulis dapat menyelesaikan penulisan skripsi ini yang berjudul “ Perbandingan Algoritma *Support Vector Machine* dan *Naïve Bayes Classifier* Pada Analisis Sentimen Aplikasi Tiktok Menggunakan Data Twitter”.

Penulisan Skripsi ini disusun untuk memenuhi salah satu syarat dalam menyelesaikan pendidikan pada Program Studi Statistika (S1) Departemen Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Padang. Dalam menyelesaikan Skripsi ini, tidak terlepas dari bimbingan dan bantuan dari berbagai pihak. Penulis menyadari bahwa skripsi ini memiliki keterbatasan sesuai dengan ilmu yang dimiliki, oleh sebab itu penulis menerima saran dan kritikan dari pembaca demi kesempurnaan isi skripsi ini penulis mengucapkan terima kasih kepada:

1. Bapak Dr. Syafriandi., M.Si., Dosen Pembimbing Skripsi.
2. Ibu Dra. Nonong Amalita., M.Si., dan Ibu Admi Salma., S.Pd, M.Si Dosen penguji.
3. Ibu Dr. Yenni Kurniawati., M.Si., Kepala Departemen Statistika dan Koordinator Program Studi S1 Statistika FMIPA UNP.
4. Bapak dan Ibu Dosen serta Tenaga Kependidikan Departemen Statistika FMIPA UNP.
5. Kedua orang tua dan saudara/i yang telah memberikan semangat, nasehat, dukungan dan bantuan moril maupun materi kepada penulis.

6. Sahabat-sahabat saya Ovel dan Dedek, terima kasih sudah menjadi teman terbaik selama menempuh perkuliahan ini dan mengajarkan banyak hal. Pengalaman yang luar biasa bersama kalian akan menjadi momen yang tidak terlupakan dan sangat dirindukan. Semoga persahabatan kita terus berlanjut dan sukses untuk kita semua.

Penulis berharap dengan disusunnya skripsi ini dapat bermanfaat bukan hanya bagi penulis sebagai syarat kelulusan, namun dapat bermanfaat bagi masyarakat terutama yang bergelut dalam pembelajaran Statistika dan Data Mining.

Padang, November 2023

Penulis

Denia Putri Fajrina

## DAFTAR ISI

ABSTRAK.....	i
KATA PENGANTAR .....	iii
DAFTAR ISI.....	iv
DAFTAR TABEL.....	vi
DAFTAR GAMBAR.....	viii
DAFTAR LAMPIRAN.....	ix
BAB I PENDAHULUAN.....	1
A. Latar Belakang Masalah.....	1
B. Batasan Masalah.....	6
C. Rumusan Masalah .....	6
D. Tujuan Penelitian.....	6
E. Manfaat Penelitian.....	7
BAB II KERANGKA TEORITIS .....	8
A. TikTok.....	8
B. Twitter .....	9
C. <i>Text Mining</i> .....	10
D. Pembobotan.....	12
E. Analisis Sentimen.....	14
F. Machine Learning .....	15
G. Klasifikasi.....	17
1. <i>Support Vector Machine</i> .....	18

2. <i>Naïve Bayes Classifier</i> .....	35
H. <i>Confusion Matrix</i> .....	39
I. <i>Word Cloud</i> .....	41
BAB III METODOLOGI PENELITIAN .....	43
A. Jenis dan Sumber Penelitian.....	43
B. Variabel Penelitian .....	43
C. Struktur Data .....	43
D. Metode Pengumpulan Data .....	45
E. Teknik Analisis Data.....	45
BAB IV HASIL DAN PEMBAHASAN .....	50
A. Hasil Penelitian .....	50
B. Pembahasan.....	74
BAB V PENUTUP .....	76
A. Kesimpulan.....	76
B. Saran.....	76
DAFTAR PUSTAKA .....	77
LAMPIRAN.....	81

## DAFTAR TABEL

<b>Tabel</b>	<b>Halaman</b>
1. Ketentuan Pelabelan VADER.....	15
2. Jenis Kernel yang Umum digunakan.....	35
3. <i>Confusion Matrix</i> .....	39
4. Kriteria Akurasi .....	40
5. Kriteria Presisi .....	41
6. Kriteria Sensitivitas .....	41
7. Variabel Penelitian.....	43
8. Struktur Data Penelitian Sebelum <i>Preprocessing</i> .....	44
9. Struktur Data <i>Preprocessing</i> .....	44
10. Contoh Ulasan Twitter.....	50
11. Contoh <i>Cleaning Data</i> .....	51
12. Contoh <i>Case Folding</i> .....	52
13. Contoh <i>Stopword</i> .....	53
14. Contoh <i>Normalization</i> .....	54
15. Contoh <i>Tokenization</i> dan <i>Stemming</i> .....	54
16. Contoh Hasil Pelabelan Data Menggunakan VADER .....	56
17. Contoh Perhitungan TF-IDF .....	57
18. Pembagian Data Set.....	58
19. Hasil Perhitungan Nilai $\alpha$ .....	60
20. Nilai Bias setiap Kernel.....	61

21. Model <i>Hyperplane</i> SVM .....	61
22. Hasil Perhitungan <i>Testing</i> SVM Kernel Linier .....	62
23. Confusion Matrix Kernel Linier .....	62
24. Hasil Perhitungan <i>Testing</i> SVM Kernel Polinomial .....	64
25. <i>Confusion Matrix</i> Kernel Polinomial.....	64
26. Hasil Evaluasi Masing-Masing Kernel.....	65
27. Frekuensi <i>Term</i> pada Setiap Kelas Sentimen .....	66
28. Hasil Perhitungan <i>Testing</i> NBC .....	68
29. <i>Confusion Matrix</i> NBC.....	69
30. Evaluasi Model NBC.....	70
31. Perbandingan Akurasi Algoritma .....	71

## DAFTAR GAMBAR

<b>Gambar</b>	<b>Halaman</b>
1. Negara-Negara Pengguna TikTok Terbesar di Dunia Awal Tahun 2023.....	2
2. Contoh Data dapat Dipisahkan Secara Linier .....	20
3. <i>Separating Hyperplanes</i> dan <i>Margin Of SVM</i> .....	21
4. <i>The Non-Separable Case</i> .....	29
5. Transformasi dari Ruang <i>Input</i> ke Ruang Fitur .....	34
6. Contoh <i>Word Cloud</i> .....	42
7. <i>Flowchart</i> Analisis Data .....	48
8. <i>Flowchart</i> SVM dan NBC .....	49
9. Frekuensi Sentimen Positif dan Negatif.....	56
10. <i>Word Cloud</i> Sentimen Positif .....	72
11. <i>Word Cloud</i> Sentimen Positif .....	73

## DAFTAR LAMPIRAN

Lampiran	Halaman
1. <i>Crawling Data</i> .....	81
2. <i>Package yang digunakan</i> .....	81
3. <i>Cleaning Data</i> .....	82
4. <i>Case Folding</i> .....	83
5. <i>Stopword</i> .....	84
6. <i>Tokenization</i> .....	85
7. <i>Normalization</i> .....	85
8. <i>Stemming</i> .....	86
9. Analisis Sentimen .....	87
10. <i>Compound Score</i> .....	87
11. <i>Filter Sentimen Netral</i> .....	89
12. Pembagian Data Set .....	89
13. TF-IDF .....	89
14. Klasifikasi SVM.....	90
15. Klasifikasi NBC .....	90
16. <i>Word Cloud</i> .....	91
17. Data Sentimen Aplikasi TikTok.....	93
18. Hasil Perhitungan <i>Testing</i> SVM Kernel Linear .....	97
19. Hasil Perhitungan <i>Testing</i> SVM Kernel Polinomial .....	99
20. Hasil Perhitungan <i>Testing</i> NBC .....	101

# BAB I

## PENDAHULUAN

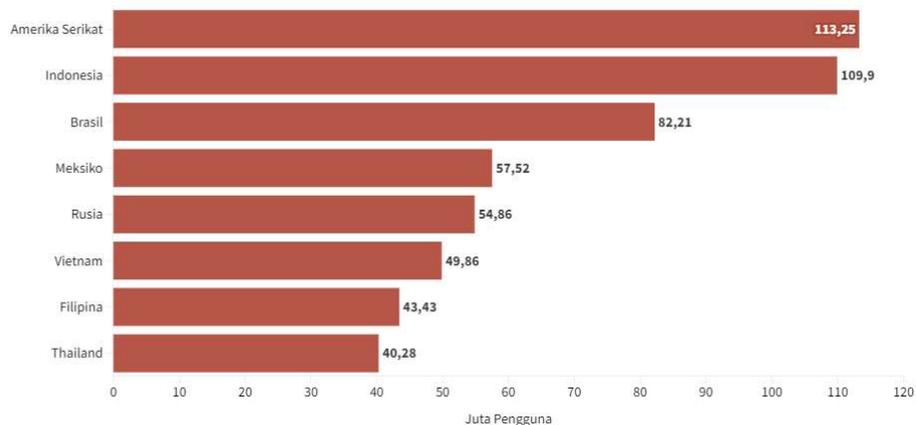
### **A. Latar Belakang Masalah**

Sejalan dengan pesatnya perkembangan teknologi, banyak jenis media yang dapat dimanfaatkan oleh manusia untuk berkomunikasi, termasuk media sosial. Melalui media sosial, seseorang dapat berbagi berita, foto, dan video dengan pengguna lainnya. Selain itu, media sosial juga berfungsi sebagai *platform* yang dapat digunakan untuk berkomunikasi dan berbagi informasi berbagai kalangan masyarakat. Saat ini, salah satu media sosial yang sangat populer adalah TikTok.

TikTok adalah aplikasi jejaring sosial dan *platform* video musik yang memungkinkan pengguna untuk membuat dan berbagi klip video pendek dengan filter, dukungan musik, efek, stiker, dan juga memiliki kolom komentar. Aplikasi ini juga memungkinkan pengguna untuk mengedit suara dalam video dan menambahkan musik ke dalamnya. Media sosial TikTok sangat banyak penggunanya. Menurut Zulqornain & Adikara (2021) penggunaan media sosial yang paling banyak diunduh adalah TikTok sejumlah 45,8 juta kali dan jumlah tersebut mengalahkan media sosial lainnya seperti WhatsApp, Youtube, Facebook dan bahkan Instagram.

TikTok menjadi aplikasi dengan pengguna terbanyak di dunia karena popularitas di kalangan generasi muda dengan berfokus pada fitur-fitur yang unik, algoritma rekomendasi konten yang canggih mendapatkan popularitas di banyak negara

Amerika Serikat, Indonesia, Brazil, Meksiko dan negara lainnya seperti disajikan pada Gambar 1.



**Gambar 1.** Negara-Negara Pengguna TikTok Terbesar di Dunia Awal Tahun 2023

(DataIndonesia.id)

Berdasarkan Gambar 1 terlihat bahwa pengguna media sosial TikTok di Indonesia mencapai 109,9 juta berada pada posisi kedua setelah Amerika Serikat yang mencapai 113,25 juta pengguna aktif pada tahun 2023. Aplikasi ini sangat sukses menarik perhatian banyak orang. Menurut laporan *Business of Apps* pada tahun 2021 pengguna TikTok di seluruh dunia didominasi oleh kelompok usia 20-29 tahun. Namun, seperti aplikasi sosial media lainnya, TikTok juga menerima banyak kritikan dari pengguna yang merasa tidak puas dengan fitur dan konten yang ditampilkan. Kritikan yang diberikan pada media sosial ini sangat menjadi perhatian apakah dampak yang dihadirkan positif maupun negatif. Komentar dan pendapat terutama yang terdapat di media sosial, merupakan sumber data yang dapat digunakan untuk mengukur popularitas suatu program atau produk yang diluncurkan. Menurut Himawan (2018) dukungan atau penolakan terhadap program

dapat dihitung berdasarkan komentar dan opini publik di media sosial dalam meningkatkan kualitas suatu produk tersebut.

Berbagai sumber komentar atau ulasan yang dapat digunakan untuk dijadikan sebagai sumber informasi seperti, Google Play Store, Sensor Tower, Twitter, dan media sosial lainnya. Twitter adalah salah satu sumber data yang dapat digunakan sebagai sumber informasi dan sering digunakan dalam analisis sentimen. Menurut Wang (2022) informasi tersebut didapat berdasarkan *tweet* dari pengguna Twitter yang juga ikut dalam menanggapi aplikasi TikTok. Menurut Kim,dkk (2006) *mention* di media sosial Twitter dapat disimpulkan dan diklasifikasikan berdasarkan kriteria yang telah ditentukan. Klasifikasi teks adalah sebuah proses yang bertujuan untuk menentukan kelas atau kategori dari suatu teks. Menurut Maarif (2016) teks ini biasa berupa kata, frasa, kalimat, paragraf, atau bahkan dokumen teks. Analisis yang dapat digunakan untuk mengklasifikasikan teks tersebut adalah analisis sentimen.

Menurut Hu & Liu (2004: 10-13) analisis sentimen secara luas mengacu pada pemrosesan bahasa alami, linguistik komputasi, dan penambangan teks yang digunakan untuk menganalisis perasaan, opini, sikap, penilaian, evaluasi, dan perasaan seseorang tentang topik tertentu seperti produk, layanan, organisasi, atau aktivitas tertentu yang digunakan. Sebelum melakukan analisis sentimen, diperlukan *preprocessing* data dengan metode *text mining* untuk mengolah data teks agar siap untuk dianalisis. Menurut Indriyani (2022) *preprocessing* data mempunyai tahapan yaitu *cleaning*, *case folding*, *tokenizing*, *normalization*, *stopword removal*, dan *stemming*. Setelah data siap digunakan, dilakukan klasifikasi ke dalam kategori

sentimen positif, negatif, maupun netral. Ada Beberapa algoritma yang dapat digunakan untuk mengklasifikasikan sentimen yaitu, *Naïve Bayes Classifier* (NBC), *Support Vector Machine* (SVM), dan *Logistic Regression*.

Penelitian ini menggunakan dua algoritma yang populer pada analisis sentimen yaitu *Support Vector Machine* (SVM) dan *Naïve Bayes Classifier* (NBC). Keunggulan pada kedua algoritma ini yaitu, pada algoritma SVM dan NBC mampu memproses data dengan sampel yang berukuran kecil. Kedua algoritma ini mampu bekerja pada dataset dengan jumlah sampel yang sedikit dan tetap menghasilkan hasil yang akurat. Pada penelitian sebelumnya yang dilakukan oleh Kristiyanti (2019) menggunakan sampel sebesar 200 *tweet* dan menunjukkan bahwa kedua algoritma tersebut memiliki akurasi baik. Algoritma ini juga dapat digunakan untuk analisis sentimen *real-time*. Menurut Permatasari dan Irhamah (2022) melakukan analisis sentimen pada ulasan Twitter dimana data *real-time* nya yaitu penggunaan *trend* vaksin *Covid-19* di masa pandemi. Kedua algoritma ini juga dapat digunakan untuk memproses data dengan dimensi yang lebih tinggi, berdasarkan penelitian Anjasmoros (2020) diperoleh bahwa semakin tinggi dimensi yang digunakan pada kedua algoritma menghasilkan *performance* data semakin baik.

Membandingkan keakuratan analisis sentimen menggunakan algoritma SVM dan NBC pada data media sosial Twitter yang dilakukan oleh Silitonga,dkk (2023) tiga kali pengujian dengan jumlah data yang berbeda, mendapatkan hasil bahwa algoritma SVM bekerja lebih baik dari pada algoritma NBC dalam hal akurasi. Studi lain Rahmadani,dkk (2022) mengusulkan untuk menerapkan algoritma NBC untuk menganalisis sentimen TikTok di media sosial menggunakan data Twitter.

Studi tersebut menemukan bahwa algoritma NBC bekerja lebih baik dalam analisis sentimen aplikasi TikTok dengan data Twitter. Perbandingan kedua algoritma ini berdasarkan penelitian sebelumnya dilihat dari keakuratan algoritma SVM dan NBC pada analisis sentimen menggunakan data media sosial, termasuk Twitter. Pada penelitian tersebut juga memberikan hasil bahwa keakuratan yang digunakan pada algoritma ini bergantung pada sumber data dan topik yang digunakan.

Pada penelitian ini juga tidak menggunakan sentimen netral yang mana pada algoritma NBC mengasumsikan bahwa fitur (kata) dalam dokumen independen satu sama lain, dan sentimen dari dokumen dihitung dengan mengalikan probabilitas setiap fitur dalam dokumen. Menurut Firdausi, dkk (2022) menjelaskan bahwa jika terdapat fitur (kata) yang memiliki unsur netral karena tidak memiliki korelasi antara sentimen positif dan negatif dapat mengurangi akurasi klasifikasi NBC. Pada algoritma SVM bekerja dengan membangun model yang memisahkan dokumen dalam kelas positif dan negatif. Menurut Firdausi, dkk (2022) juga menjelaskan sentimen netral tidak memiliki korelasi dengan sentimen positif dan negatif, ini dapat membingungkan model SVM dan mengurangi akurasi.

Berdasarkan latar belakang yang telah diuraikan sebelumnya, dilakukan penelitian terhadap pengguna (*user*) *Twitter* terkait topik tentang aplikasi TikTok menggunakan algoritma *Naïve Bayes Classifier* dan *Support Vector Machine*, dengan menjadikan topik tersebut sebagai bahan kajian yang tertuang dalam penelitian dengan judul “*Perbandingan Metode Support Machine Vector Dan Naïve Bayes Classifier Pada Analisis Sentimen Aplikasi Tiktok Menggunakan Data Twitter*”.

## **B. Batasan Masalah**

Batasan masalah dalam penelitian ini adalah.

1. Data yang digunakan merupakan data ulasan mengenai aplikasi *TikTok* di Twitter bulan Desember 2022 hingga bulan Januari 2023 sebanyak 1836 ulasan.
2. Ulasan yang diambil merupakan ulasan yang berbahasa Indonesia
3. Klasifikasi sentimen *tweet* dibagi menjadi dua, yaitu sentimen positif dan sentimen negatif

## **C. Rumusan Masalah**

Rumusan masalah dalam penelitian adalah.

1. Diantara algoritma *Support Vector Machine* dan *Naïve Bayes Classifier* manakah yang memiliki akurasi tertinggi dalam melakukan analisis sentimen pada Aplikasi TikTok di Twitter?
2. Visualisasi kata apa saja yang sering muncul pada sentimen di ulasan Aplikasi TikTok di Twitter ?

## **D. Tujuan Penelitian**

Tujuan penelitian ini adalah sebagai berikut .

1. Untuk mengetahui algoritma terbaik antara *Support Vector Machine* dan *Naïve Bayes Classifier* dalam melakukan analisis sentimen pada aplikasi TikTok di Twitter.
2. Untuk mengetahui penciri sentimen positif dan negatif pada ulasan Aplikasi TikTok di Twitter yang telah dilakukan klasifikasi.

### **E. Manfaat Penelitian**

Manfaat pada penelitian ini adalah sebagai berikut.

1. Bagi penulis, dapat menambah pengetahuan tentang algoritma *Support Vector Machine* dan *Naïve Bayes Classifier* dalam melakukan klasifikasi sentimen pada aplikasi TikTok di media sosial Twitter.
2. Bagi pengembang aplikasi TikTok, dapat menjadi bahan evaluasi untuk meningkatkan fitur-fitur pada produknya.
3. Bagi masyarakat, hasil yang diperoleh dapat sebagai gambaran umum mengenai persepsi masyarakat terhadap penggunaan aplikasi TikTok berdasarkan data Twitter.