

**ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP KINERJA
KEMENTERIAN KEUANGAN MENGGUNAKAN ALGORITMA NAÏVE
BAYES**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



diajukan oleh

FARHAN RALDIFAN 19.11.3244

Kepada

PROGRAM SARJANA

PROGRAM STUDI INFORMATIKA

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2023

**ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP KINERJA
KEMENTERIAN KEUANGAN MENGGUNAKAN ALGORITMA NAÏVE
BAYES**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



diajukan oleh

FARHAN RALDIFAN 19.11.3244

Kepada

PROGRAM SARJANA

PROGRAM STUDI INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS AMIKOM YOGYAKARTA

YOGYAKARTA

HALAMAN PERSETUJUAN

HALAMAN PERSETUJUAN

SKRIPSI

ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP KINERJA
KEMENTERIAN KEUANGAN MENGGUNAKAN ALGORITMA NAÏVE

BAYES

yang disusun dan diajukan oleh

Farhan Raaidiffa

19.11.3244

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 18 Agustus 2023

Dosen Pembimbing,



Remini, M.Koni,

NIK. 190302246

HALAMAN PENGESAHAN SKRIPSI

HALAMAN PENGESAHAN SKRIPSI

ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP KINERJA KEMENTERIAN KEUANGAN MENGGUNAKAN ALGORITMA NAÏVE BAYES

yang disusun dan diajukan oleh

FARHAN RALDIFAN

19.11.3244

Telah dipertahankan di depan Dewan Penguji
pada tanggal 18 Agustus 2023

Susunan Dewan Penguji

Nama Penguji

Anna Bafta, M.Kom
NIK. 190302190

Nurhikmah, M.Kom
NIK. 190302345

Ramini, M.Kom
NIK. 190302246

Tanda Tangan



Skrripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 18 Agustus 2023

DEKAN FAKULTAS ILMU KOMPUTER



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

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Farhan Raldifan

NIM : 19.11.3244

Menyatakan bahwa Skripsi dengan judul berikut:

ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP KINERJA KEMENTERIAN KEUANGAN MENGGUNAKAN ALGORITMA NAÏVE BAYES

Dosen Pembimbing : Rumiati, 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 Agustus 2023

Yang Menyatakan,



Farhan Raldifan

HALAMAN PERSEMBAHAN

Puji syukur kehadirat Allah SWT, atas limpahan rahmat dan karunia-Nya sehingga saya dapat menyelesaikan laporan skripsi ini. Hasil karya ini saya persembahkan kepada diri saya sendiri yang telah berjuang semaksimal mungkin di bangku perkuliahan dan kepada orang-orang yang telah berjasa memberi dukungan tiada henti :

1. Kedua orang tua saya yang senantiasa mendoakan setiap langkah serta berbagai macam dukungan untuk mengantarkan saya menuju kesuksesan.
2. Seluruh keluarga besar saya yang senantiasa mendoakan, mendukung, dan menantikan kelulusan saya.
3. Dosen pembimbing saya, Rumini, M.Kom, yang senantiasa memberikan arahan, perbaikan selama penyusunan dan pelaksanaan penelitian.
4. Orang-orang terkasih dan para sahabat terdekat saya yang tidak dapat saya sebutkan satu persatu yang telah selalu memberikan dukungan, menghibur dan menemani.
5. Teman-teman mahasiswa Informatika yang tidak dapat saya sebutkan satu persatu,

KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Allah SWT yang telah memberikan rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan penyusunan skripsi dengan judul "Analisis Sentimen Pengguna Twitter Terhadap Kinerja Kementerian Keuangan Menggunakan Algoritma Naive Bayes". Skripsi ini disusun sebagai salah satu syarat untuk menyelesaikan studi di Program Sarjana Universitas Amikom Yogyakarta.

Dengan selesainya skripsi ini, maka penulis tidak lupa mengucapkan terimakasih kepada :

1. Bapak Prof. Dr. M. Suyanto, MM selaku Rektor Universitas Amikom Yogyakarta.
2. Hanif Al Fatta, S.Kom., M.Kom, Selaku Dekan Fakultas Ilmu Komputer.
3. Rumini, M.Kom selaku dosen pembimbing saya yang telah banyak memberikan pengarahan bagi penulis dalam pembuatan skripsi.
4. Bapak-Ibu Dosen Universitas Amikom Yogyakarta, atas ilmu dan pengetahuan yang telah diberikan kepada penulis selama duduk di bangku kuliah.
5. Kedua orang tua saya, Bapak dan Almarhumah Ibu yang tak pernah berhenti untuk selalu memberikan doa.

Penulis tentunya menyadari bahwa pembuatan skripsi ini masih banyak sekali kekurangan-kekurangan dan kelemahan-kelemahannya. Oleh karena itu penulis berharap kepada semua pihak agar dapat menyampaikan kritik dan saran yang membangun untuk menambah kesempurnaan skripsi ini. Namun penulis tetap berharap skripsi ini akan bermanfaat bagi semua pihak yang membacanya.

Yogyakarta, 18 Agustus 2023

Farhan Raldifan

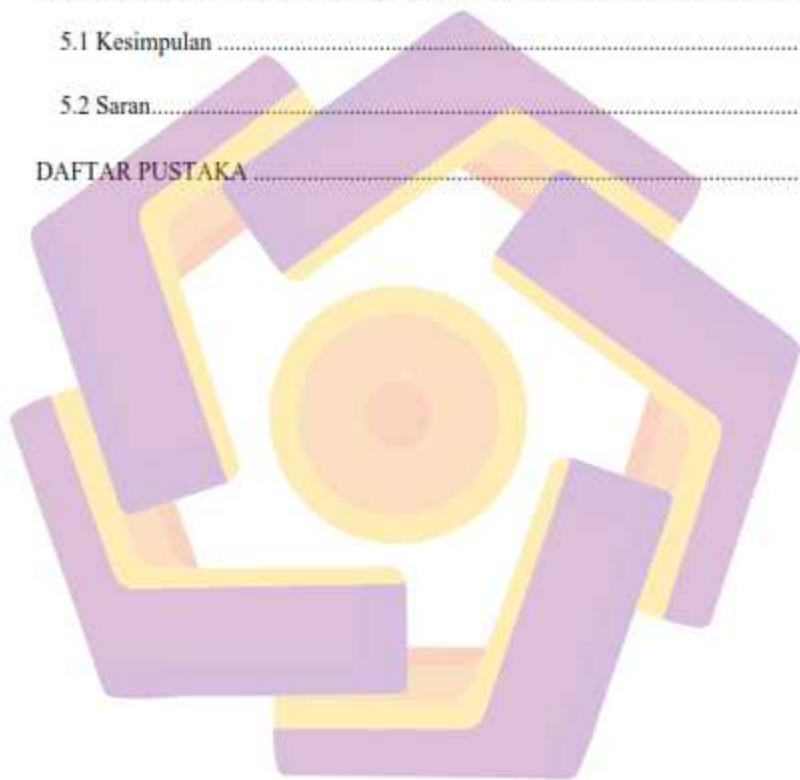
DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN SKRIPSI.....	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI.....	iv
HALAMAN PERSEMBAHAN.....	v
KATA PENGANTAR.....	vi
DAFTAR ISI.....	vii
DAFTAR TABEL.....	xi
DAFTAR GAMBAR.....	xii
DAFTAR LAMBANG DAN SINGKATAN.....	xiv
DAFTAR ISTILAH.....	xv
INTISARI.....	xvi
Abstract.....	xvii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan masalah.....	3
1.3 Batasan Masalah.....	3
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	4
1.6 Sistematika Penulisan.....	4
BAB II TINJAUAN PUSTAKA.....	6
2.1 Studi Literatur.....	6

2.2 Dasar Teori.....	13
2.2.1 Twitter.....	13
2.2.2 <i>Python</i>	13
2.2.3 Sentimen Analisis.....	13
2.2.4 Text Mining.....	14
2.2.5 <i>Pre-processing</i>	14
2.2.6 <i>Balancing Data</i>	16
2.2.7 <i>Lexicon</i>	16
2.2.8 Klasifikasi.....	17
2.2.9 <i>Multinomial Naïve Bayes</i>	17
2.2.10 <i>Confussion Matrix</i>	18
2.2.11 <i>Wordcloud</i>	19
BAB III METODE PENELITIAN.....	20
3.1 Objek Penelitian.....	20
3.2 Alur Penelitian.....	21
3.2.1 Pengumpulan Data.....	21
3.2.2 <i>Preprocessing Data</i>	21
3.2.3 <i>Labelling Dataset</i>	23
3.2.4 Pembagian data uji dan data latih.....	23
3.2.5 Pembobotan kata.....	23

3.2.6 SMOTE.....	23
3.2.7 Metode Klasifikasi.....	24
3.2.8 Evaluasi Model.....	24
3.3 Alat dan Bahan.....	24
BAB IV HASIL DAN PEMBAHASAN.....	26
4.1 Pengumpulan Data.....	26
4.2 <i>Preprocessing Data</i>	27
4.2.1 <i>Case Folding</i>	27
4.2.2 <i>Cleansing Data</i>	28
4.2.3 <i>Tokenizing</i>	29
4.2.4 <i>Normalization</i>	29
4.2.5 <i>Stopword Removal</i>	30
4.2.6 <i>Stemming</i>	31
4.3 <i>Labelling Dataset</i>	32
4.4 <i>Wordcloud</i>	34
4.5 <i>Splitting Data</i>	37
4.6 <i>Term Frequency–Inverse Document Frequency (TF-IDF)</i>	37
4.7 SMOTE.....	38
4.8 <i>Multinomial Naïve Bayes</i>	38
4.9 <i>Diskusi Hasil</i>	39

4.9.1 Naïve Bayes 10% data uji 90% data latih	39
4.9.2 Naïve Bayes 20% data uji 80% data latih	40
4.9.3 Naïve Bayes 30% data uji 70% data latih	40
BAB V KESIMPULAN DAN SARAN	42
5.1 Kesimpulan	42
5.2 Saran.....	42
DAFTAR PUSTAKA	43



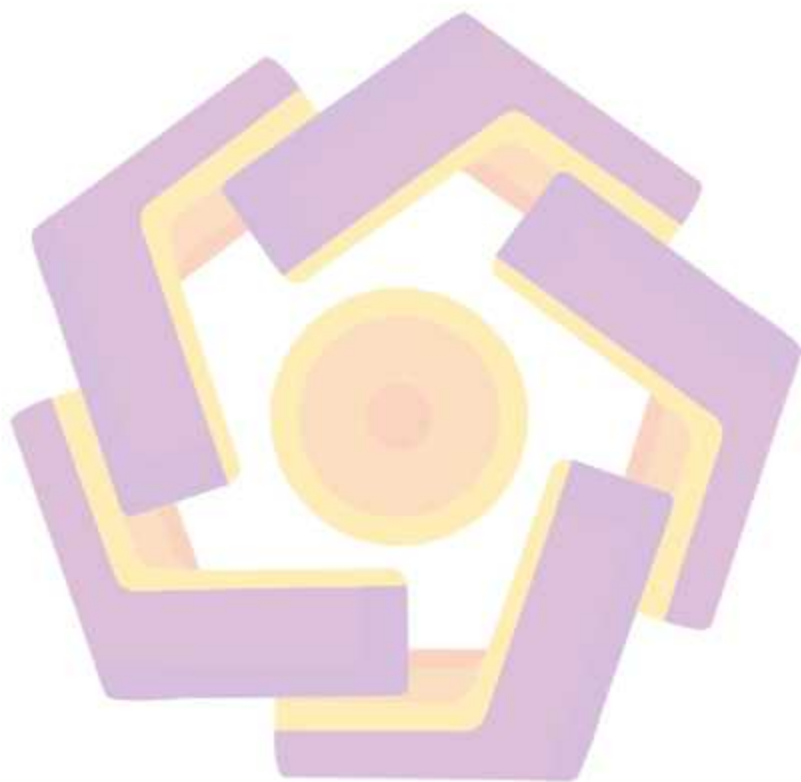
DAFTAR TABEL

Tabel 2.1 Keaslian Penelitian	9
Tabel 2.2 Confusion Matrix	18
Tabel 4.1 Case Folding.....	28
Tabel 4.2 Cleansing Data	29
Tabel 4.3 Tokenizing.....	30
Tabel 4.4 Normalization.....	30
Tabel 4.5 Stopword Removal.....	31
Tabel 4.6 Stemming	32
Tabel 4.7 Labelling Data.....	34
Tabel 4.8 TF-IDF	38
Tabel 4.9 Akurasi	41

DAFTAR GAMBAR

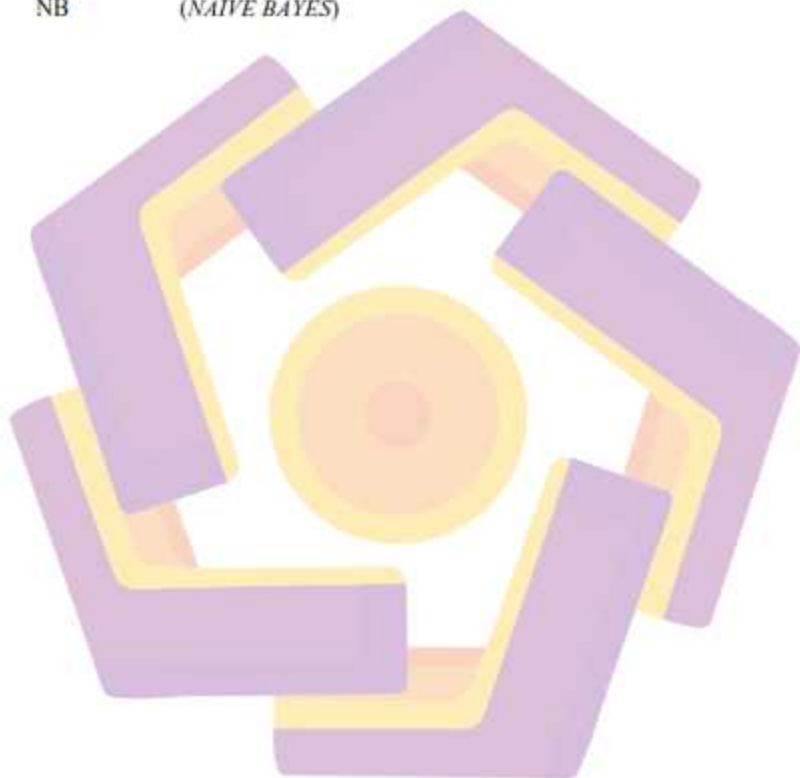
Gambar 3.1 Flowchart Penelitian	20
Gambar 3.2 Flowchart Pre-processing	21
Gambar 4.1 Crawling data	26
Gambar 4.2 Hasil Crawling Data	27
Gambar 4.3 Case Folding	27
Gambar 4.4 Cleansing data	28
Gambar 4.5 Duplicate Removal	29
Gambar 4.6 Tokenizing	29
Gambar 4.7 Normalization	30
Gambar 4.8 Stopword Removal	31
Gambar 4.9 Stemming	32
Gambar 4.10 Lexicon	33
Gambar 4.11 komposisi data setelah labelling	33
Gambar 4.12 Positive wordcloud	34
Gambar 4.13 Script Wordcloud	35
Gambar 4.14 Negative Wordcloud	36
Gambar 4.15 Script Negative Wordcloud	36
Gambar 4.16 Splitting data 10% data uji 90% data latih	37
Gambar 4.17 Splitting data 20% data uji 80% data latih	37
Gambar 4.18 Splitting data 30% data uji 70% data latih	37
Gambar 4.19 TF-IDF	37
Gambar 4.20 Sebelum SMOTE	38
Gambar 4.21 Setelah SMOTE	38
Gambar 4.22 Multinomial Naïve Bayes	38
Gambar 4.23 Confusion Matrix	39
Gambar 4.24 Print Classification Report	39
Gambar 4.25 Classification Report perbandingan rasio pertama	39
Gambar 4.26 Classification Report perbandingan rasio kedua	40

Gambar 4.27 Classification Report perbandingan rasio ketiga 41



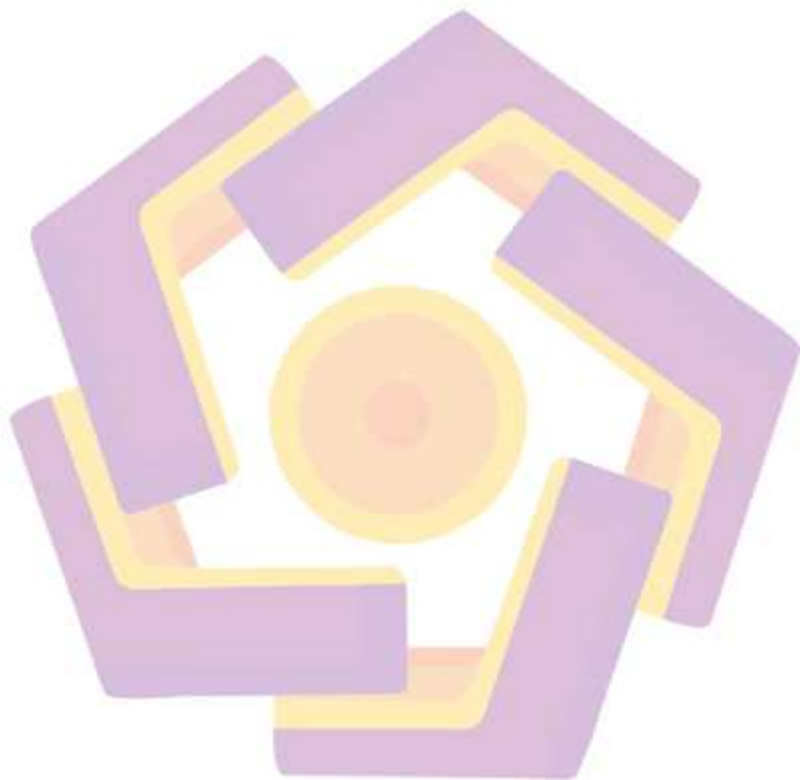
DAFTAR LAMBANG DAN SINGKATAN

TFIDF	(Term Frequency Inverse Document Frequency)
SMOTE	(Synthetic Minority Oversampling Technique)
NB	(<i>NAÏVE BAYES</i>)



DAFTAR ISTILAH

Vector Besaran yang mempunyai arah



INTISARI

Awal tahun ini khususnya pada bulan Februari hingga akhir Mei, Kementerian Keuangan mendapat banyak sorotan dari publik atas kegiatannya dibalik layar yang hampir tak terlihat oleh masyarakat awam. Dampak dari beberapa kejadian tersebut dapat membuat masyarakat tidak percaya lagi terhadap jajaran petinggi ataupun pejabat pemerintahan kita. Dari permasalahan tersebut saya ingin mengetahui bagaimana tanggapan masyarakat terhadap kinerja pemerintah khususnya kementerian keuangan. Pada penelitian ini media yang saya gunakan untuk mengukur tanggapan masyarakat ialah Twitter, karena pada media sosial tersebut sering digunakan untuk menyampaikan beberapa pendapat dan argumen perseorangan maupun organisasi. Twitter sendiri merupakan media sosial berbasis teks, gambar dan video. Namun beberapa pengguna twitter lebih banyak menggunakan media sosial ini dengan mengkritik maupun menyampaikan pendapatnya terhadap isu yang sedang dibicarakan. Pada penelitian ini saya menilai argumen-argumen tersebut menggunakan analisis sentiment, klasifikasi lexicon based dan algoritma Naïve Bayes. Lexicon based sendiri merupakan analisis teks digital untuk menentukan apakah teks tersebut bernilai positif atau negatif. Namun sebelum Lexicon Based menilai, teks yang telah terambil harus di-preprocessing terlebih dahulu. Serta pada penelitian ini menggunakan balancing data SMOTE. Naïve Bayes merupakan salah satu metode yang dapat digunakan untuk mengklasifikasikan dan memprediksi data. Hasil dari algoritma Naïve Bayes pada penelitian ini telah menampilkan akurasi mencapai 76,42%.

Kata kunci: Analisis sentimen, Naïve bayes, twitter, kemenkeu, lexicon based.

Abstract

The beginning of this year, especially from February to the end of May, the Ministry of Finance received a lot of attention from the public for its behind-the-scenes activities which were almost invisible to the general public. The impact of some of these events can make people no longer trust our high-ranking ranks or government officials. From these problems I want to know how the public's response to the government's performance, especially the ministry of finance. In this study, the medium that I used to measure public response was Twitter, because social media is often used to convey several opinions and arguments from individuals and organizations. Twitter itself is a social media based on text, images and videos. However, some Twitter users use this social media more by criticizing or expressing their opinions on the issues being discussed. In this research, I assess these arguments using sentiment analysis, lexicon based classification and the Naïve Bayes algorithm. Lexicon based itself is a digital text analysis to determine whether the text is positive or negative. But before Lexicon Based judges, the text that has been taken must be pre-processed first. As well as in this study using SMOTE data balancing. Naïve Bayes is a method that can be used to classify and predict data. The results of the Naïve Bayes algorithm in this study have shown an accuracy of up to 76.42%.

Keyword: Sentiment Analysis, Naïve Bayes, Twitter, The Ministry of Finance, Lexicon Based.