

**KLASIFIKASI TREN BISNIS BERDASARKAN SENTIMEN
PENGGUNA TWITTER DENGAN *MACHINE LEARNING***

SKRIPSI

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



disusun oleh

YUNITA NURISFA MAYA SAKTI

19.11.3079

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2023

**KLASIFIKASI TREN BISNIS BERDASARKAN SENTIMEN
PENGGUNA TWITTER DENGAN *MACHINE LEARNING***

SKRIPSI

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



disusun oleh

YUNITA NURISFA MAYA SAKTI

19.11.3079

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2023

HALAMAN PERSETUJUAN

SKRIPSI

**KLASIFIKASI TREN BISNIS BERDASARKAN SENTIMEN PENGGUNA
TWITTER DENGAN MACHINE LEARNING**

yang disusun dan diajukan oleh

Yunita Nurisfa Maya Sakti

19.11.3079

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 5 September 2023

Dosen Pembimbing,



Windha Mega Pradnya D. M. Kom

NIK. 190302185

HALAMAN PENGESAHAN

SKRIPSI

**KLASIFIKASI TREN BISNIS BERDASARKAN SENTIMEN PENGGUNA
TWITTER DENGAN *MACHINE LEARNING***

yang disusun dan diajukan oleh

Yunita Nurisfa Maya Sakti

19.11.3079

Telah dipertahankan di depan Dewan Penguji
pada tanggal 15 Agustus 2023

Susunan Dewan Penguji

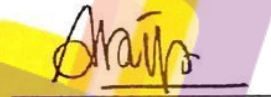
Nama Penguji

Tanda Tangan

Asro Nasiri, Drs. M.Kom
NIK. 190302152



Arif Dwi Laksito, M.Kom
NIK. 190302150



Windha Mega Pradnya D. M.Kom
NIK. 190302185



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 15 Agustus 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 : Yunita Nurisfa Maya Sakti
NIM : 19.11.3079

Menyatakan bahwa Skripsi dengan judul berikut:

KLASIFIKASI TREN BISNIS BERDASARKAN SENTIMEN PENGGUNA TWITTER DENGAN MACHINE LEARNING

Dosen Pembimbing : Windha Mega Pradnya D, 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, 15 Agustus 2023

Yang Mervatakan,



Yunita Nurisfa Maya Sakti

HALAMAN PERSEMBAHAN

Segala puji bagi Allah SWT dengan kemurahan dan ridho-Nya, sehingga skripsi ini dapat ditulis dengan baik dan lancar hingga selesai. Penulisan skripsi ini sungguh tidak akan mungkin dapat terselesaikan tanpa adanya bantuan dan dukungan dari segala pihak. Peneliti ingin mengucapkan banyak rasa terima kasih dan mempersembahkan skripsi kepada :

1. Diri saya sendiri yang telah berjuang dan bertahan untuk menyelesaikan penulisan skripsi ini dengan sangat baik.
2. Kedua orang tua dan seluruh anggota keluarga yang selalu memberi bantuan, dukungan, dan nasehat sehingga penulis dapat menyelesaikan skripsi dengan sangat baik.
3. Ibu Windha Mega Pradnya D, M.Kom selaku dosen pembimbing yang selalu memberi saran dan masukan dalam penulisan skripsi ini.
4. Fadillah dan Siti selaku teman dekat penulis dan rekan seperjuangan penulis di AMIKOM yang selalu memberi saran dan menemani penulis selama proses penulisan skripsi.
5. PANAROMA (Paul, Nabila, Rony, Salma) yang selalu memberi semangat, dan hiburan kepada penulis melalui karya-karya dan konten yang diberikan kepada penulis.
6. Teman-teman kelas 19 Informatika 08 selaku rekan seperjuangan penulis selama menempuh pendidikan.
7. Semua pihak yang berperan dalam proses penulisan skripsi ini yang tidak bias disebutkan satu persatu

KATA PENGANTAR

Puji Syukur bagi Allah SWT atas rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan penyusunan skripsi dengan judul “KLASIFIKASI TREN BISNIS BERDASARKAN SENTIMEN PENGGUNA TWITTER DENGAN MACHINE LEARNING” ini dengan baik. Selama proses penulisan skripsi ini, penulis banyak mendapatkan bantuan, pengarahan, dan bimbingan dari berbagai pihak. Oleh karena itu, pada kesempatan ini sebagai wujud rasa hormat, penulis mengucapkan terima kasih yang sebesar besarnya :

1. Prof. Dr. Suyanto, M.M selaku Rektor Universitas Amikom Yogyakarta.
2. Ibu Windha Mega Pradnya D, M.Kom selaku Ketua Program Studi S1 Informatika Universitas Amikom Yogyakarta dan juga selaku dosen pembimbing yang selalu memberi saran dan masukan dalam penulisan skripsi ini.
3. Bapak Arif Dwi Laksito, M.Kom dan Bapak Asro Nasiri, Drs. M.Kom selaku dosen penguji yang telah memberi saran dan masukan dalam penulisan skripsi ini.
4. Kedua orangtua, adik, dan keluarga penulis yang senantiasa mendoakan dan memberikan dukungan kepada penulis selama penulisan skripsi ini.
5. Semua pihak yang senantiasa menjadi teman dan penghibur penulis sehingga skripsi ini dapat selesai dengan baik.

Penulis menyadari bahwa skripsi ini masih terdapat banyak kekurangan baik dalam isi maupun cara penyajian. Oleh karena itu, penulis mengharapkan kritik dan saran guna perbaikan di masa mendatang.

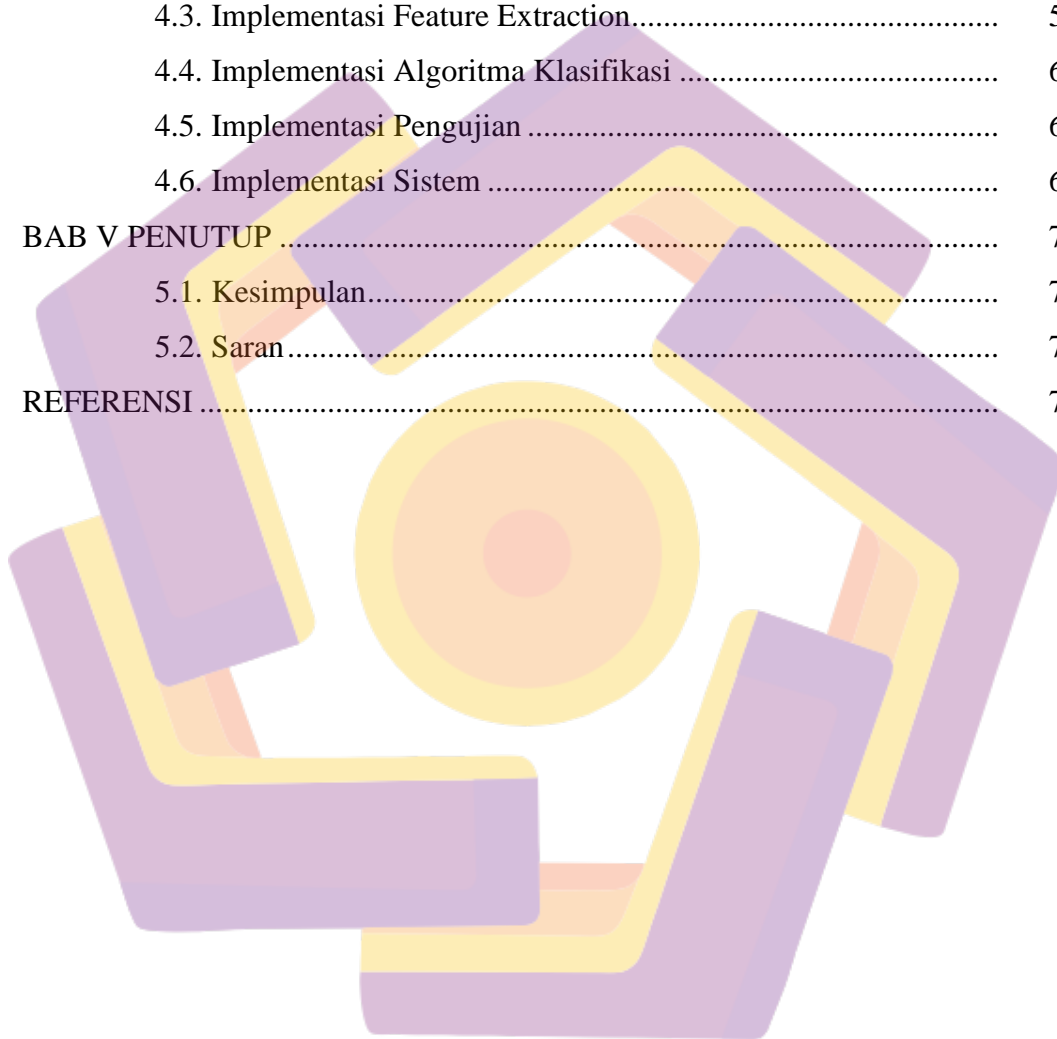
Yogyakarta, 15 Agustus 2023

Penulis

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI.....	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI.....	vii
DAFTAR TABEL.....	ix
DAFTAR GAMBAR	xi
INTISARI	xii
ABSTRACT.....	xiii
BAB 1 PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	2
1.3. Batasan Masalah.....	3
1.4. Tujuan Penelitian.....	3
1.5. Manfaat Penelitian.....	3
1.6. Sistematika Penulisan.....	4
BAB II LANDASAN TEORI.....	6
2.1. Studi Literatur.....	6
2.2. Dasar Teori	13
BAB III METODE PENELITIAN	25
3.1. Gambaran Umum	25
3.2. Alur Penelitian.....	26
3.3. Alat dan Bahan	35

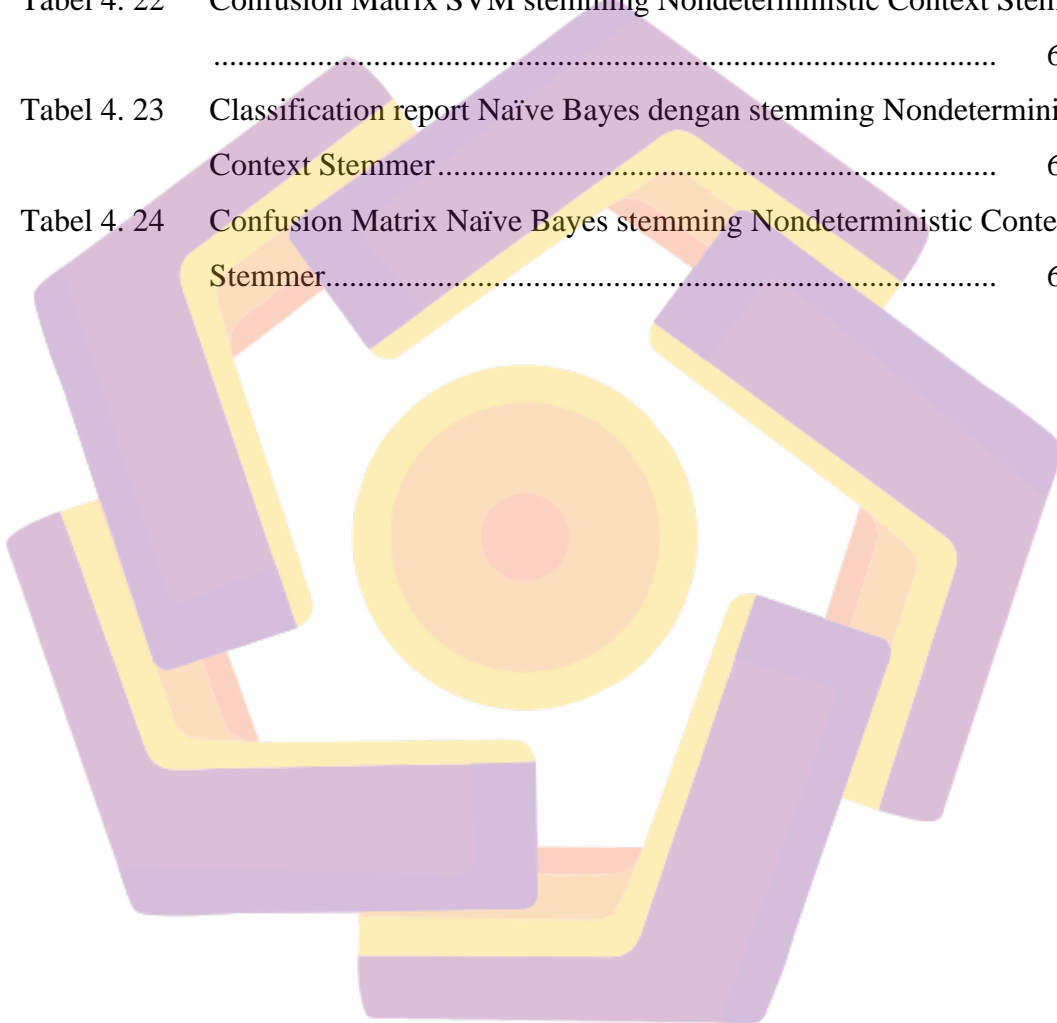
3.4. Perancangan Interface	36
BAB IV HASIL DAN PEMBAHASAN	39
4.1. Implementasi Pengumpulan Dataset	39
4.2. Implementasi Preprocessing Data	45
4.3. Implementasi Feature Extraction.....	59
4.4. Implementasi Algoritma Klasifikasi	60
4.5. Implementasi Pengujian	61
4.6. Implementasi Sistem	68
BAB V PENUTUP	72
5.1. Kesimpulan.....	72
5.2. Saran.....	73
REFERENSI	74



DAFTAR TABEL

Tabel 2.1	Keaslian Penelitian	9
Tabel 2.2	Confusion Matrix	22
Tabel 3. 1	Tabel label data	28
Tabel 3.2	Tabel Kebutuhan perangkat keras dan spesifikasinya.....	35
Tabel 3.3	Tabel Kebutuhan perangkat lunak.....	36
Tabel 4.1	Tabel Raw Data fashion	39
Tabel 4.2	Tabel Raw Data kuliner.....	40
Tabel 4.3	Tabel Raw Data jasa.....	42
Tabel 4.4	Tabel Raw Data properti	43
Tabel 4.5	Tabel Raw Data kerajinan Tangan	44
Tabel 4.6	Tabel hasil filtering data.....	46
Tabel 4. 7	Tabel hasil Case Folding dan Data Cleaning	47
Tabel 4.8	Tabel hasil Normalization	49
Tabel 4.9	Tabel hasil Normalization	51
Tabel 4.10	Tabel hasil Stopword Removal	52
Tabel 4.11	Tabel hasil Stemming Sastrawi	54
Tabel 4. 12	Tabel frekuensi <i>unique word</i> dari data stemming Sastrawi	56
Tabel 4.13	Tabel hasil Stemming Nondeterministic Context Stemmer.....	57
Tabel 4. 14	Tabel frekuensi <i>unique word</i> dari data stemming Nondeterministic Context Stemmer.....	58
Tabel 4.15	Tabel hasil evaluasi stemming Sastrawi.....	61
Tabel 4.16	Tabel hasil classification report SVM stemming Sastrawi	62
Tabel 4.17	Confusion Matrix SVM stemming Sastrawi	63
Tabel 4. 18	Tabel classification report Naïve Bayes stemming Sastrawi ...	63
Tabel 4. 19	Confusion Matrix SVM stemming Sastrawi	64

Tabel 4. 20	Tabel hasil evaluasi SVM dengan stemming Nondeterministic Context Stemmer.....	64
Tabel 4. 21	Classification report SVM stemming Nondeterministic Context Stemmer.....	65
Tabel 4. 22	Confusion Matrix SVM stemming Nondeterministic Context Stemmer	66
Tabel 4. 23	Classification report Naïve Bayes dengan stemming Nondeterministic Context Stemmer.....	66
Tabel 4. 24	Confusion Matrix Naïve Bayes stemming Nondeterministic Context Stemmer.....	67



DAFTAR GAMBAR

Gambar 2.1	Gambar konsep dasar klasifikasi SVM	20
Gambar 3.1	Gambar alur penelitian	27
Gambar 3.2	Gambar alur data filtering	31
Gambar 3.3	Gambar alur data case folding & cleansing.....	31
Gambar 3.4	Aalur data normalization	31
Gambar 3.5	Alur tokenization	32
Gambar 3.6	Alur stopword removal.....	32
Gambar 3.7	Alur data stemming	33
Gambar 3.8	Halaman Untuk Menampilakn Raw Data	36
Gambar 3.9	Halaman untuk menampilkan data yang telah dipreprocessing	37
Gambar 3.10	Halaman untuk menampilkan hasil evaluasi klasifikasi SVM.	37
Gambar 3.11	Halaman untuk menampilkan evaluasi klasifikasi Naïve Bayes	38
Gambar 3.12	Halaman untuk melakukan klasifikasi teks dengan SVM.....	38
Gambar 3.13	Halaman untuk melakukan klasifikasi dengan Naïve Bayes....	38
Gambar 4.1	Gambar penerapan TF-IDF	60
Gambar 4.2	Gambar penerapan split data	60
Gambar 4.3	Gambar penerapan model SVM	60
Gambar 4.4	Gambar penerapan model Naïve Bayes	60
Gambar 4.5	Gambar diagram perbandingan akurasi.....	68
Gambar 4.6	Halaman tampilan menu Raw Data.....	68
Gambar 4.7	Halaman tampilan data yang telah dipreprocessing	69
Gambar 4.8	Halaman tampilan hasil evaluasi klasifikasi SVM.....	69
Gambar 4.9	Halaman untuk tampilan evaluasi klasifikasi Naïve Bayes.....	70
Gambar 4.10	Halaman tampilan klasifikasi teks dengan SVM	70
Gambar 4.11	Halaman tampilan klasifikasi dengan Naïve Bayes	71

INTISARI

Salah satu cara untuk melakukan riset tren bisnis yaitu dengan melakukan analisis tren pada google tren. Akan tetapi, pada google tren hanya memberikan hasil tren keseluruhan baik sentiment positif maupun negatif. Untuk mengetahui sentiment tren bisnis dapat dilakukan analisis sentiment dari pengguna twitter. Analisis sentiment tersebut dapat dilakukan dengan metode machine learning. Menurut penelitian sebelumnya, diantara beberapa metode klasifikasi metode naïve bayes dan support vector machine menghasilkan akurasi yang cukup tinggi. Naïve bayes merupakan metode klasifikasi yang mampu mengestimasi data dengan baik dengan prediksi trend an pola, sedangkan support vector machine adalah metode yang implementasinya menggunakan hyperplane yang memudahkan untuk melakukan analisis sentiment. Pada penelitian ini metode support vector machine dan naïve bayes akan dibandingkan untuk melakukan analisis sentiment tren bisnis berdasarkan twitt pengguna twitter. Data yang digunakan yaitu data twitt dari tanggal 1 Januari 2022 hingga 31 Desember 2022. Sebelum dilakukan pemodelan SVM dan Naïve bayes data dilakukan preprocessing dengan beberapa tahap yaitu filtering, case folding dan data cleaning, normalization, tokenization, stopword removal dan stemming. Pada penelotian ini stemming dilakukan menggunakan library Sastrawi dan Nondeterministic Context Stemmer. Perbandingan akurasi antara metode SVM dan Naïve bayes dilakukan dengan menggunakan metode confusion matrix. Hasil pengujian yang didapatkan menghasilkan bahwa akurasi SVM lebih tinggi jika dibandingkan dengan akurasi naïve bayes, yaitu SVM mendapatkan akurasi sebesar 88% pada proses stemming menggunakan sastrawi dan 87% pada proses stemming menggunakan Nondeterministic Context Stemmer sedangkan Naïve bayes sebesar 75% pada proses stemming menggunakan Sastrawi dan Nondeterministic Context Stemmer.

Kata kunci: Tren bisnis, sentiment analisis, support vector machine, naïve bayes.

ABSTRACT

One way to research business trends is to do trend analysis on Google Trends. However, Google Trends only provides overall trend results, both positive and negative sentiments. To find out business trend sentiment, sentiment analysis from Twitter users can be carried out. Sentiment analysis can be done using machine learning methods. According to previous research, among several classification methods, naïve Bayes method and support vector machine produce a fairly high accuracy. Naïve Bayes is a classification method that is able to estimate data well by predicting trends and patterns, while the support vector machine is a method whose implementation uses a hyper line which makes it easy to do sentiment analysis. In this study, the support vector machine and naïve Bayes methods will be compared to carry out business trend sentiment analysis based on Twitter user tweets. The data used is tweet data from January 1, 2022 to December 31, 2022. Prior to SVM modeling and Naïve Bayes data preprocessing was carried out in several stages, namely filtering, case folding and data cleaning, normalization, tokenization, stopword removal and stemming. In this study, stemming was carried out using Sastrawi libraries and Nondeterministic Context Stemmer libraries. Comparison of accuracy between the SVM and Naïve Bayes methods was carried out using the confusion matrix method. The test results obtained show that SVM accuracy is higher when compared to naïve Bayes accuracy, namely SVM obtains an accuracy of 88% in the stemming process using Sastrawi libraries and 87% in the stemming process using Nondeterministic Context Stemmer while Naïve Bayes is 75% in the stemming process using Literature and Nondeterministic Context Stemmer.

Keyword: Business trend, sentiment analys, support vector machine, naïve bayes.