

**PENGUJIAN KERNEL SVM PADA SENTIMENT ANALYSIS
MENGENAI TELKOMSEL**

SKRIPSI

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



disusun oleh

PANGESTU FREMMUZAR

19.11.2955

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2023**

PENGUJIAN KERNEL SVM PADA SENTIMENT ANALYSIS

MENGENAI TELKOMSEL

HALAMAN JUDUL

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana

Program Studi Informatika



disusun oleh

PANGESTU FREMMUZAR

19.11.2955

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA**

YOGYAKARTA

2023

HALAMAN PERSETUJUAN

SKRIPSI

**PENGUJIAN KERNEL SVM PADA SENTIMENT ANALYSIS MENGENAI
TELKOMSEL**

yang disusun dan diajukan oleh

PANGESTU FREMMUZAR

19.11.2955

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 27 Februari 2023

Dosen Pembimbing,

Anna Baita M. Kom
NIK. 190302290

HALAMAN PENGESAHAN

SKRIPSI

PENGUJIAN KERNEL SVM PADA SENTIMENT ANALYSIS MENGENAI
TELKOMSEL

yang disusun dan diajukan oleh

PANGESTU FIE SEMUZAR

19.11.2955

Telah dipertahankan di depan Dewan Penguji
pada tanggal 27 Februari 2023

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Anna Haida, M.Kom
NIK. 190302290



Anggit Dwi Hartanto, M.Kom
NIK. 190302163



Arif Dwi Laksana, M.Kom
NIK. 190302150



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 27 Februari 2023

DEKAN FAKULTAS ILMU KOMPUTER



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

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : **PANGESTU FREMMUZAR**
NIM : **19.11.2955**

Menyatakan bahwa Skripsi dengan judul berikut:

PENGUJIAN KERNEL SVM PADA SENTIMENT ANALYSIS MENGENAI TELKOMSEL

Dosen Pembimbing : **Anna Baita, 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, 27 Februari 2023

Yang Menyatakan,



The image shows a rectangular stamp with a grid pattern on the left side. To the right of the stamp, there is a handwritten signature in black ink. The signature appears to be 'Pangestu Fremmuzar'. Below the signature, the name 'Pangestu Fremmuzar' is printed in a small font.

Pangestu Fremmuzar

HALAMAN PERSEMBAHAN

Puji dan Syukur atas kehadiran Allah SWT, yang telah melimpahkan rahmat, hidayah dan karunia-Nya, atas kemudahan, petunjuk dan kekuatan yang diberikan-Nya, sehingga penulis dapat menyelesaikan penyusunan skripsi yang berjudul “Analisis Sentimen Mengenai Produk Telkomsel Menggunakan Algoritma SVM”, dan diajukan sebagai salah satu syarat untuk mendapat gelar Sarjana pada jurusan Informatika, Universitas Amikom Yogyakarta. Karya tulis ini penulis persembahkan kepada :

1. Pertama saya ucapkan terima kasih kepada ayah saya Suyadi, yang sudah menyiapkan kebutuhan finansial saya dalam menempuh pendidikan hingga berhasil mencapai titik ini sebagai sarjana pertama dalam keluarga saya.
2. Kedua saya ucapkan terima kasih kepada ibu saya Mirna Erwinda, yang selalu mendukung dan menyemangati saya dalam menjalani pendidikan perguruan tinggi hingga merantau ke luar daerah. Kemudian saya ucapkan juga terima kasih kepada ibu saya yang selalu mengingatkan saya untuk beribadah dan selalu memberikan doa yang terbaik untuk saya.
3. Ketiga saya ucapkan terima kasih kepada adik saya Jelita Dwi Anggraeni yang selalu mendukung dan menyemangati saya dalam menempuh pendidikan di luar daerah.
4. Terakhir, saya ucapkan terima kasih kepada teman, saudara, dan orang yang selalu ada untuk membantu dan menyemangati saya dalam penulisan skripsi ini.

KATA PENGANTAR

Puji dan Syukur atas kehadiran Allah SWT yang telah melimpahkan rahmat, hidayah dan karunia-Nya, atas kemudahan, petunjuk dan kekuatan yang diberikannya, sehingga penulis dapat menyelesaikan penyusunan skripsi yang berjudul “Analisis Sentimen Mengenai Produk Telkomsel Menggunakan Algoritma SVM”, dan diajukan sebagai salah satu syarat untuk mendapat gelar Sarjana pada jurusan Informatika, Universitas Amikom Yogyakarta. Dalam penyusunannya, penulis memperoleh banyak bantuan dari berbagai pihak, oleh karena itu penulis mengucapkan terimakasih yang sebesar-besarnya kepada :

1. Ibu Anna Baita, M. Kom sebagai dosen pembimbing yang telah memberikan kontribusi berupa bimbingan, motivasi, kritik dan saran dalam menyelesaikan skripsi ini.
2. Semua Dosen Fakultas Ilmu Komputer Universitas Amikom Yogyakarta, yang tidak bisa penulis sebutkan satu persatu. Terimakasih atas ilmu yang sudah diajarkan selama ini.
3. Kedua orang tua dan keluarga besar yang senantiasa mendoakan dan menyemangati serta memberi dukungan.
4. Semua pihak yang telah membantu dalam penyusunan skripsi ini dari awal hingga akhir.

Tidak ada kata selain terima kasih yang sebesar-besarnya kepada semua pihak yang telah membantu saya sampai titik ini. Penulis menyadari bahwa penulisan skripsi ini masih memiliki banyak kekurangan. Untuk semua kritik dan saran yang bersifat membangun sangat penulis harapkan dan akan diterima dengan lapang dada.

Yogyakarta, 22 Februari 2023



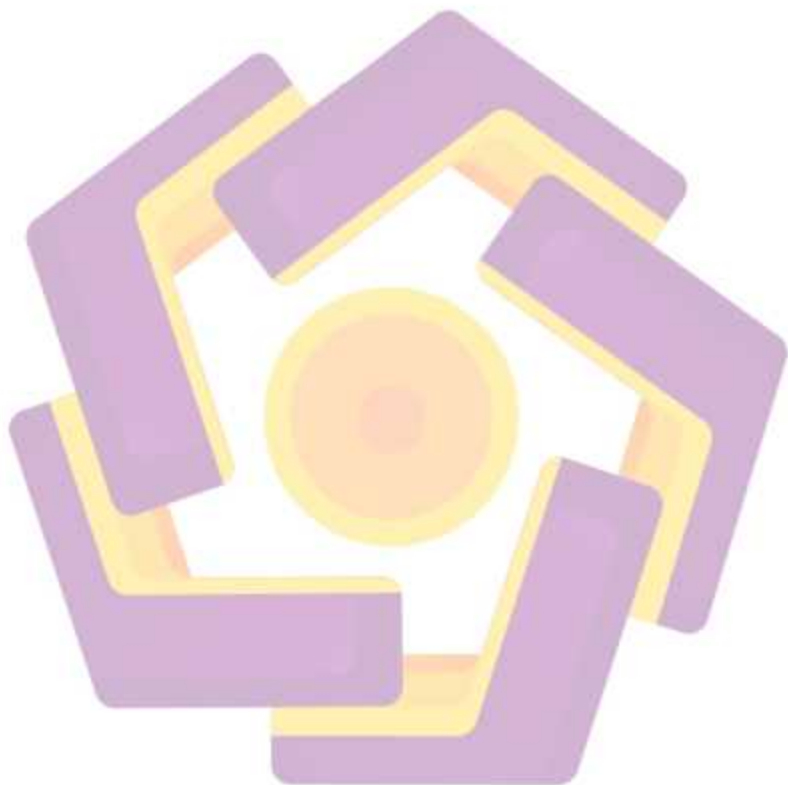
Pangestu Fremmuzar

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iii
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI	vii
DAFTAR TABEL	x
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	2
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.6 Sistematika Penulisan	4
BAB II TINJAUAN PUSTAKA	5
2.1. Studi Literatur	5
2.2. Dasar Teori	9
2.2.1 Telkomsel	9
2.2.2 Twitter	9
2.2.3 <i>Machine Learning</i>	9

2.2.4	<i>Text Mining</i>	10
2.2.5	<i>Support Vector Machine</i>	14
2.2.6	<i>Confusion Matrix</i>	19
BAB III METODE PENELITIAN		21
3.1	Objek Penelitian	21
3.2	Alur Penelitian	21
3.2.1	<i>Scraping Data</i>	21
3.2.2	<i>Dataset Preparation</i>	23
3.2.3	<i>Labeling Dataset</i>	24
3.2.4	<i>Preprocessing Data</i>	25
3.2.5	TF-IDF	34
3.2.6	<i>Splitting Data</i>	35
3.2.7	SVM Implementation	35
3.2.8	Evaluasi	36
3.3	Alat dan Bahan	36
BAB IV HASIL DAN PEMBAHASAN		38
4.1	Pengumpulan dan Pengolahan <i>Dataset</i>	38
4.2	Visualisasi Kata	39
4.3	Implementasi SVM	39
4.4	Implementasi Kernel SVM	41
4.4.1	Kernel Linear	41
4.4.2	Kernel RBF	42
4.4.3	Kernel Polynomial	43
4.4.4	Kernel Sigmoid	44
4.5	Evaluasi dan Perbandingan	45
4.5.1	Perbandingan Kernel	45
4.5.2	<i>Confusion Matrix</i> Kernel Sigmoid	46

4.5.3 Diskusi Hasil.....	47
BAB V PENUTUP	52
5.1 Kesimpulan	52
5.2 Saran	52
DAFTAR PUSTAKA	53

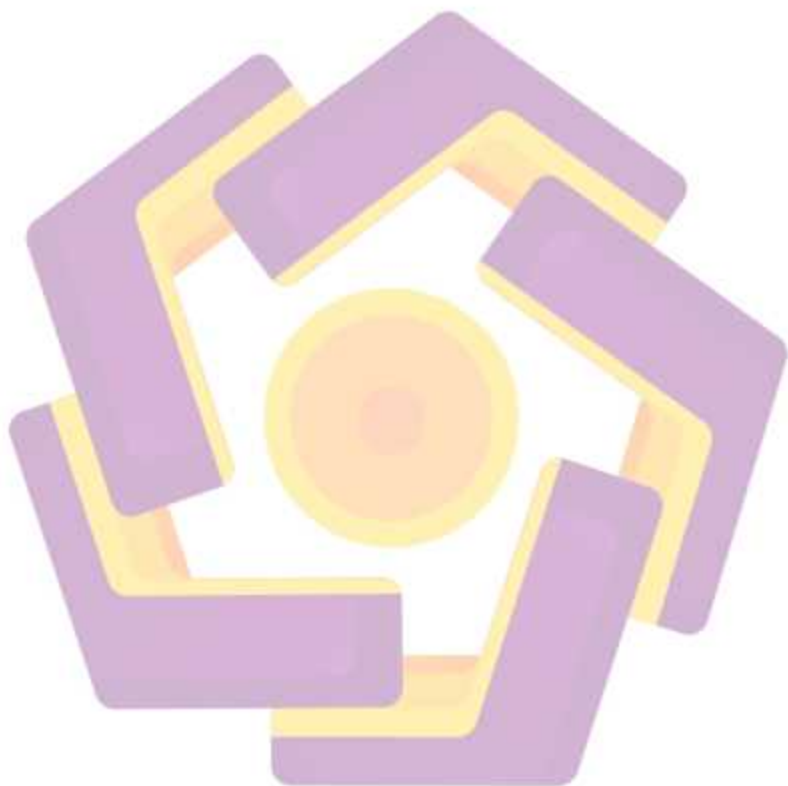


DAFTAR TABEL

Tabel 2.1 Keaslian Penelitian	7
Tabel 3.1 Sampel Dataset	22
Tabel 3.2 Contoh Dataset Berlabel	24
Tabel 3.3 Contoh Hasil Cleaning	25
Tabel 3.4 Contoh Hasil Casefolding	27
Tabel 3.5 Contoh Hasil Tokenizing	28
Tabel 3.6 Contoh Hasil Normalization	29
Tabel 3.7 Contoh Sampel Kamus Normalization	30
Tabel 3.8 Contoh Hasil Stemming	31
Tabel 3.9 Contoh Hasil Stopword Removal	32
Tabel 3.10 Contoh Hasil Detokenize	33
Tabel 3.11 Informasi Alat Penelitian	37
Tabel 3.12 Library Python Yang Digunakan	37
Tabel 4.1 Total Penghapusan Data	38
Tabel 4.2 Pembagian Jumlah Dataset	40
Tabel 4.3 Hasil Pengujian Parameter Kernel Linear	41
Tabel 4.4 Hasil Pengujian Parameter Kernel RBF	42
Tabel 4.5 Hasil Pengujian Parameter Kernel Polynomial	43
Tabel 4.6 Hasil Pengujian Parameter Kernel Sigmoid	44
Tabel 4.7 Perbandingan Hasil Tiap Kernel	45
Tabel 4.8 Kata Yang Sering Muncul Pada Sentimen Positif	49

Tabel 4.9 Kata Yang Sering Muncul Pada Sentimen Negatif 50

Tabel 4.10 Perbandingan Hasil Klasifikasi 51

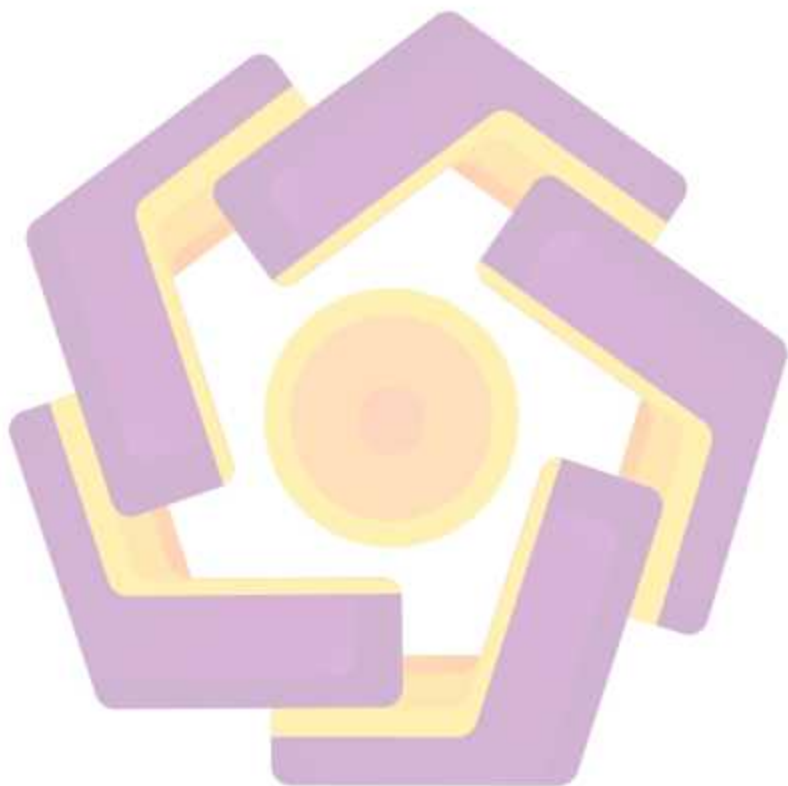


DAFTAR GAMBAR

Gambar 2.1 Ilustrasi Hyperplane Terbaik dan Margin Maksimal	15
Gambar 2.2 Ilustrasi Kernel Linear	16
Gambar 2.3 Ilustrasi Kernel RBF	16
Gambar 2.4 Ilustrasi Kernel Sigmoid	17
Gambar 2.5 Ilustrasi Kernel Polynomial	18
Gambar 2.6 Confusion Matrix 2x2	19
Gambar 3.1 Alur Penelitian	21
Gambar 3.2 Alur Scraping Data	21
Gambar 3.3 Alur Dataset Preparation	23
Gambar 3.4 Alur Preprocessing Data	25
Gambar 3.5 Alur Proses TF-IDF	34
Gambar 3.6 Hasil TF-IDF	34
Gambar 3.7 Alur Proses Implementasi SVM	35
Gambar 4.1 Sentimen Yang Sering Muncul	39
Gambar 4.2 Perbandingan Performa Pembagian Dataset	40
Gambar 4.3 Performa SVM Default	41
Gambar 4.4 Perbandingan Performa Kernel	46
Gambar 4.5 Confusion Matrix Kernel Sigmoid	46
Gambar 4.6 Perhitungan Performa Kernel Sigmoid	47
Gambar 4.7 Confusion Matrix Klasifikasi Kernel Sigmoid	48
Gambar 4.8 Perhitungan Performa Klasifikasi Kernel Sigmoid	48

Gambar 4.9 Sentimen Positif Yang Sering Muncul 49

Gambar 4.10 Sentimen Negatif Yang Sering Muncul 50



DAFTAR LAMBANG DAN SINGKATAN



SVM	Support Vector Machines
NB	Naive Bayes
EM	Maximum Entropy
DT	Decision Tree
KNN	K-Nearest Neighbors
RF	Random Forest
ANN	Artificial Neural Network
LR	Linear Regression
PT	Perseroan Terbatas
GSM	Global System for Mobile Communications
TF	Term Frequency
IDF	Inverse Document Frequency
RBF	Radial Basic Function
TP	True Positif
TN	True Negatif
FP	False Positif
FN	False Negatif
KBBI	Kamus Besar Bahasa Indonesia

DAFTAR ISTILAH

Scraping	Proses pengumpulan data dengan cara
Provider	Penyedia jasa layanan internet
Sentimen	Pendapat atau pandangan
Netizen	Warganet atau warga internet
Preprocessing	Pengolahan atau pemrosesan data
Binary-class	Sebuah kelas yang terdiri dari 2 jenis kelas
Multi-class	Sebuah kelas yang terdiri lebih dari 2 jenis kelas



INTISARI

Interconnection-networking (Internet) adalah sistem jaringan komputer di seluruh dunia yang menghubungkan satu sama lain dalam segala hal dunia menggunakan *standart internet protocol suite*. Telkomsel merupakan penyedia layanan internet di Indonesia yang diluncurkan pada tanggal 26 Mei 1995. Sebagai penjedia jasa layanan internet dengan pengguna terbanyak, Telkomsel menjadi pusat perhatian para pengguna internet di Indonesia. Hal ini mengundang opini dan sudut pandang pengguna terhadap Telkomsel yang biasa disebut dengan sentimen. Salah satu media yang biasa digunakan untuk mengutarakan sebuah opini dan sudut pandang adalah Twitter. Twitter merupakan platform media sosial yang sering menjadi tempat untuk berbagi dan menyebarkan berita, diskusi ide, dan opini pengguna Twitter. Dalam penelitian ini, algoritma yang digunakan adalah *Support Vector Machine*. Pada algoritma *Support Vector Machine*, terdapat kernel trick yang bisa digunakan untuk mendapatkan hasil yang lebih optimal. Kernel yang digunakan yaitu linear, rbf, polynomial, dan sigmoid. Jumlah sentimen yang dianalisis sebanyak 537 cuitan dan sudah melewati tahap *preprocessing*. Sebuah sentimen diklasifikasi menjadi 2 label, yaitu positif dan negatif. Berdasarkan hasil pengujian, kernel sigmoid memiliki performa terbaik dengan nilai *accuracy* sebesar 0.950, *precision* sebesar 0.945, *recall* sebesar 0.860, dan *f1-score* sebesar 0.896.

Kata kunci: telkomsel, svm, preprocessing, sentimen, kernel.

ABSTRACT

Interconnection-networking (Internet) is a worldwide system of computer networks that connect in all respects of the world using the standard Internet protocol suite. Telkomsel is an internet service provider in Indonesia which was launched on May 26, 1995. As an internet service provider with the most users, Telkomsel has become the center of attention of internet users in Indonesia. This invites user opinions and perspectives on Telkomsel, which is commonly referred to as sentiment. One of the media commonly used to express an opinion and point of view is Twitter. Twitter is a social media platform that is often a place for sharing and spreading the news, and discussing the ideas, and opinions of Twitter users. In this research, the algorithm used is Support Vector Machine. In the Support Vector Machine algorithm, there is a kernel trick that can be used to get more optimal results. The kernels used are linear, RBF, polynomial, and sigmoid. The number of sentiments analyzed was 337 tweets and they had already passed the preprocessing stage. A sentiment is classified into 2 labels, namely positive and negative. Based on the test results, the sigmoid kernel has the best performance with an accuracy value of 0.950, a precision of 0.945, a recall of 0.860, and an f1-score of 0.896.

Keyword: telkomsel, svm, preprocessing, sentiment, kernel.

