



**ANALISIS SENTIMEN TERHADAP LAYANAN PROVIDER
TELKOMSEL PADA JEJARING SOSIAL *TWITTER*
MENGUNAKAN ALGORITMA *SUPPORT VECTOR MACHINE***

SKRIPSI

ROIHAN MUFLI ARJUNA

1710511037

UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAKARTA

FAKULTAS ILMU KOMPUTER

PROGRAM STUDI INFORMATIKA

2021



**ANALISIS SENTIMEN TERHADAP LAYANAN PROVIDER
TELKOMSEL PADA JEJARING SOSIAL *TWITTER*
MENGUNAKAN ALGORITMA *SUPPORT VECTOR MACHINE***

SKRIPSI

Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh Gelar

Sarjana Komputer

ROIHAN MUFLI ARJUNA

1710511037

UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAKARTA

FAKULTAS ILMU KOMPUTER

PROGRAM STUDI INFORMATIKA

2021

PERNYATAAN ORISINALITAS

Tugas skripsi ini adalah hasil karya sendiri, dan semua sumber yang dikutip maupun yang dirujuk telah saya nyatakan dengan benar.

Nama : Roihan Mufli Arjuna

NIM : 1710511037

Tanggal : 24 Juni 2021

Bilamana dikemudian hari ditemukan ketidaksesuaian dengan pernyataan saya ini, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan yang berlaku.

Jakarta, 24 Juni 2021

Yang Menyatakan,



(Roihan Mufli Arjuna)

PERNYATAAN PERSETUJUAN PUBLIKASI

Sebagai civitas akademik Universitas Pembangunan Nasional Veteran Jakarta, saya yang bertandatangan di bawah ini:

Nama : Roihan Mufli Arjuna

NIM : 1710511037

Fakultas : Ilmu Komputer

Program Studi : S1 Informatika

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Pembangunan Nasional Veteran Jakarta Hak Bebas Royalti Non eksklusif (*Non-Exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul:

**ANALISIS SENTIMEN TERHADAP LAYANAN PROVIDER
TELKOMSEL PADA JEJARING SOSIAL *TWITTER* MENGGUNAKAN
ALGORITMA *SUPPORT VECTOR MACHINE***

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti ini Universitas Pembangunan Nasional Veteran Jakarta berhak menyimpan, mengalih media/formatkan, mengelola dalam bentuk pangkalan kata (Basis data), merawat dan mempublikasikan Skripsi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta. Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta

Pada Tanggal : 24 Juni 2021

Yang Menyatakan,



(Roihan Mufli Arjuna)

LEMBAR PENGESAHAN

Dengan ini dinyatakan bahwa Skripsi berikut :

Nama : Roihan Mufli Arjuna

NIM : 1710511037

Program Studi : Informatika

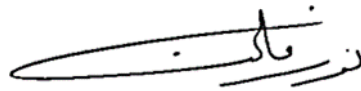
Judul Skripsi : Analisis Sentimen Terhadap Layanan Provider Telkomsel
Pada Jejaring Sosial *Twitter* Menggunakan Algoritma
Support Vector Machine

Telah berhasil dipertahankan di hadapan Tim Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Komputer pada Program Studi S1 Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta.



Iin Ernawati, S.Kom., M.Si

Penguji I



Noor Falih, S.Kom., M.T.

Penguji II



Yuni Widiastiwi, S.Kom., Msi.

Pembimbing I



Nurul Chamidah, S.Kom., M.Kom

Pembimbing II



Dr. Ernawati, M.Kom.

Dekan



Yuni Widiastiwi, S.Kom., Msi.

Ketua Program Studi

Ditandatangani : Jakarta

Tanggal Pengesahan : 16 Juli 2021



ANALISIS SENTIMEN TERHADAP LAYANAN PROVIDER TELKOMSEL PADA JEJARING SOSIAL TWITTER MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE

Roihan Mufli Arjuna

Abstrak

Twitter merupakan salah satu jejaring sosial yang digunakan untuk menampung keluhan, kritik, serta saran dari pelanggan oleh banyak perusahaan besar salah satunya perusahaan telekomunikasi yang ada di Indonesia yaitu PT Telekomunikasi Selular atau biasa disebut Telkomsel. Pada penelitian ini yang akan diteliti adalah sentimen pelanggan berdasarkan dari cuitan twitter yang ditandai dengan akun Telkomsel dengan username @Telkomsel. Data yang diambil adalah data *tweet*, data tersebut diambil dengan menggunakan Twitter API (*Application Programming Interface*). Kemudian data *tweet* tersebut diberi label positif dan negatif yang akan diberikan oleh anotator. Setelah itu, dilakukan praproses seperti pembersihan data, *case folding*, *normalization*, *stopword removal*, *stemming*, dan *tokenization* kemudian pembobotan kata dengan *Term Frequency – Inverse Document Frequency*. Sebelum proses klasifikasi, dilakukan *oversampling* data dengan teknik SMOTE karena terdapat 624 positif dan 2214 negatif yang berarti data tidak seimbang. Setelah dilakukan *oversampling*, data menjadi seimbang yaitu 2214 positif dan 2214 negatif. Lalu data dibagi 70% data latih dan 30% data uji secara acak. Algoritma yang digunakan untuk melakukan klasifikasi adalah *support vector machine*. Hasil klasifikasi yang didapatkan adalah akurasi sebesar 0.93, *precision* sebesar 0.91, *recall* sebesar 0.96, *specificity* sebesar 0.91, *F1 Score* sebesar 0.93 .

Kata Kunci : Telkomsel, Analisis Sentimen, Klasifikasi, *Support Vector Machine*

SENTIMENT ANALYSIS OF PROVIDER SERVICE TELKOMSEL ON SOCIAL NETWORKING TWITTER USING SUPPORT VECTOR MACHINE ALGORITHM

Roihan Mufli Arjuna

Abstract

Twitter is one of social media is used to have complains, critics, and suggestions from customer by established companies. One of them is Telecommunication Company in Indonesia, it is PT Telekomunikasi Selular or so-called Telkomsel. This research will study sentiment analysis customer on Twitter by tagging Telkomsel account with username @Telkomsel. Fetched data is data tweet, they were fetched by using Twitter API (Application Programming Interface). Then data tweet was labeled positive and negative by annotator. After that, preprocessing data such as cleaning, case folding, normalization, stopword removal, stemming, and tokenization then term weighting with Term Frequency – Inverse Document Frequency (TF-IDF). Before classification process, data is oversampled by using SMOTE technique because there is 614 positive and 2214 negative which mean it is imbalanced data. After oversampling, data become balanced it is 2214 positive and 2214 negative. Then data was divided to 70% as data training and 30% as data testing randomly. Algorithm is used to classify is support vector machine. Classification result is accuracy has 0.93, precision has 0.91, recall has 0.96, specificity has 0.91, F1 score has 0.93.

Kata Kunci : Telkomsel, Sentiment Analysis, Classification, Support Vector Machine

KATA PENGANTAR

Puji syukur sebesar besarnya penulis panjatkan kehadirat Allah SWT atas segala karunia dan ridha-Nya, sehingga Skripsi dengan judul “Analisis Sentimen Terhadap Layanan Provider Telkomsel Pada Jejaring Sosial Twitter Menggunakan Algoritma *Support Vector Machine*” berhasil diselesaikan. Tak lupa, penulis ingin mengucapkan banyak terima kasih kepada:

1. Orang tua, keluarga yang selalu memberikan dukungan kepada penulis sehingga dapat menyelesaikan skripsi ini.
2. Ibu Yuni Widiastiwi, S.Kom., Msi. dan Ibu Nurul Chamidah, S.Kom., M.Kom selaku dosen pembimbing yang telah memberikan saran yang bermanfaat selama proses pembuatan Proposal hingga menyelesaikan skripsi.
3. Bapak Noor Falih, S.Kom., M.T selaku dosen pembimbing akademik.
4. Ibu Yuni Widiastiwi, S.Kom., Msi. selaku Kaprodi Informatika yang telah memberikan informasi mengenai tugas akhir.
5. Ibu Dr. Ermatita, M.Kom. selaku dekan Fakultas Ilmu Komputer.
6. Teman-teman Informatika 2017 yang telah berjuang bersama dalam setiap proses perkuliahan serta saling memberikan semangat untuk dapat menyelesaikan Skripsi.
7. Teman-teman Badan Eksekutif Mahasiswa FIK Periode 2019/2020
8. Teman-teman KSM Android UPNVJ Periode 2019/2020 dan Periode 2020/2021.

Dan semua pihak yang telah membantu penulis dalam menyelesaikan skripsi ini.

Jakarta, 24 Juni 2021

Daftar Isi

COVER	i
JUDUL	ii
PERNYATAAN ORISINALITAS	iii
PERNYATAAN PERSETUJUAN PUBLIKASI	iv
LEMBAR PENGESAHAN	v
Abstrak	vi
Abstract	vii
KATA PENGANTAR	viii
Daftar Isi.....	ix
Daftar Tabel	xii
Daftar Gambar.....	xiii
Daftar Simbol	xiv
BAB 1 PENDAHULUAN	Error! Bookmark not defined.
1.1 Latar Belakang	Error! Bookmark not defined.
1.2 Rumusan Masalah	Error! Bookmark not defined.
1.3 Tujuan Penelitian.....	Error! Bookmark not defined.
1.4 Manfaat Penelitian.....	Error! Bookmark not defined.
1.5 Ruang Lingkup	Error! Bookmark not defined.
1.6 Luaran Penelitian.....	Error! Bookmark not defined.
1.7 Sistematika Penulisan.....	Error! Bookmark not defined.
BAB 2 TINJAUAN PUSTAKA	Error! Bookmark not defined.
2.1 Twitter	Error! Bookmark not defined.
2.2 Text Mining	Error! Bookmark not defined.
2.3 Analisis Sentimen.....	Error! Bookmark not defined.

2.4	Praproses	Error! Bookmark not defined.
2.4.1	Pembersihan Data.....	Error! Bookmark not defined.
2.4.2	Case Folding	Error! Bookmark not defined.
2.4.3	Normalization.....	Error! Bookmark not defined.
2.4.4	Stopword Removal.....	Error! Bookmark not defined.
2.4.5	Stemming	Error! Bookmark not defined.
2.4.6	Tokenization.....	Error! Bookmark not defined.
2.5	Term Frequency – Inverse Document Frequency (TF – IDF)	Error! Bookmark not defined.
2.6	Klasifikasi.....	Error! Bookmark not defined.
2.6.1	Support Vector Machine	Error! Bookmark not defined.
2.6.2	Evaluasi	Error! Bookmark not defined.
2.7	Kappa Statistic.....	Error! Bookmark not defined.
2.8	Review Riset yang Relevan.....	Error! Bookmark not defined.
BAB 3 METODOLOGI PENELITIAN.....		Error! Bookmark not defined.
3.1	Identifikasi Masalah	Error! Bookmark not defined.
3.2	Studi Pustaka	Error! Bookmark not defined.
3.3	Pengumpulan Data	Error! Bookmark not defined.
3.4	Pelabelan Data	Error! Bookmark not defined.
3.5	Praproses	Error! Bookmark not defined.
3.5.1	Pembersihan data	Error! Bookmark not defined.
3.5.2	Case Folding	Error! Bookmark not defined.
3.5.3	Normalization.....	Error! Bookmark not defined.
3.5.4	Stopword Removal.....	Error! Bookmark not defined.
3.5.5	Stemming	Error! Bookmark not defined.
3.5.6	Tokenization.....	Error! Bookmark not defined.

3.6	Pembobotan dengan TF-IDF	Error! Bookmark not defined.
3.7	SMOTE (Synthetic Minority Oversampling Technique)	Error! Bookmark not defined.
	Bookmark not defined.	
3.8	Klasifikasi.....	Error! Bookmark not defined.
3.9	Evaluasi	Error! Bookmark not defined.
3.10	Alat yang Digunakan.....	Error! Bookmark not defined.
3.11	Jadwal Penelitian	Error! Bookmark not defined.
BAB 4 HASIL DAN PEMBAHASAN.....		Error! Bookmark not defined.
4.1.	Data	Error! Bookmark not defined.
4.2.	Pelabelan Data	Error! Bookmark not defined.
4.3.	Praproses	Error! Bookmark not defined.
4.3.1.	Pembersihan Data.....	Error! Bookmark not defined.
4.3.2.	Case Folding	Error! Bookmark not defined.
4.3.3.	Normalization.....	Error! Bookmark not defined.
4.3.4.	Stopword Removal.....	Error! Bookmark not defined.
4.3.5.	Stemming	Error! Bookmark not defined.
4.3.6.	Tokenization.....	Error! Bookmark not defined.
4.4.	Pembobotan dengan TF-IDF	Error! Bookmark not defined.
4.5.	SMOTE (Synthetic Minority Oversampling Technique)	Error! Bookmark not defined.
	Bookmark not defined.	
4.6.	Klasifikasi.....	Error! Bookmark not defined.
4.7.	Evaluasi	Error! Bookmark not defined.
BAB 5 PENUTUP		Error! Bookmark not defined.
5.1.	Kesimpulan.....	Error! Bookmark not defined.
5.2.	Saran.....	Error! Bookmark not defined.
Daftar Pustaka		Error! Bookmark not defined.

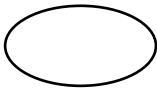



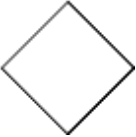
Daftar Tabel

Tabel 2.1 Confusion Matrix	Error! Bookmark not defined.
Tabel 2.2 Variasi Antar-Pengamat	Error! Bookmark not defined.
Tabel 2.3 Interpretasi Kappa	Error! Bookmark not defined.
Tabel 3.1 Jadwal Penelitian.....	Error! Bookmark not defined.
Tabel 4.1 Hasil Anotasi.....	Error! Bookmark not defined.
Tabel 4.2 Data Sampel Pembersihan Data	Error! Bookmark not defined.
Tabel 4.3 Data Sampel Case Folding	Error! Bookmark not defined.
Tabel 4.4 Sampel Kata-Kata Normalisasi	Error! Bookmark not defined.
Tabel 4.5 Data Sampel Normalization	Error! Bookmark not defined.
Tabel 4.6 Data Sampel Stopword Removal	Error! Bookmark not defined.
Tabel 4.7 Data Sampel Stemming.....	Error! Bookmark not defined.
Tabel 4.8 Data Sampel Tokenization	Error! Bookmark not defined.
Tabel 4.9 Nilai TF, IDF, dan TF-IDF Sampel	Error! Bookmark not defined.
Tabel 4.10 Sampel Data	Error! Bookmark not defined.
Table 4.11 Contoh Hasil Setelah Oversampling Dengan SMOTE	Error! Bookmark not defined.
Tabel 4.12 Proses Grid Search CV	Error! Bookmark not defined.
Tabel 4.13 Confusion Matrix Hasil Klasifikasi	Error! Bookmark not defined.

Daftar Gambar

Gambar 3.1 Tahapan Penelitian	Error! Bookmark not defined.
Gambar 4.1 Hasil Pelabelan Sentimen.....	Error! Bookmark not defined.
Gambar 4.2 Proses Pembersihan Data	Error! Bookmark not defined.
Gambar 4.3 Proses Case Folding	Error! Bookmark not defined.
Gambar 4.4 Proses Normalization	Error! Bookmark not defined.
Gambar 4.5 Proses Stopword Removal	Error! Bookmark not defined.
Gambar 4.6 Proses Stemming	Error! Bookmark not defined.
Gambar 4.7 Proses Tokenization	Error! Bookmark not defined.
Gambar 4.8 Kode Program Grid Search CV.....	Error! Bookmark not defined.
Gambar 4.9 Hasil Grid Search CV.....	Error! Bookmark not defined.

Daftar Simbol

Simbol Flowchart			
NO	GAMBAR	NAMA	KETERANGAN
1		<i>Terminal (start, end)</i>	Menggambarkan bagaimana kegiatan dimulai atau kegiatan berakhir.
2		<i>Flow Direction</i>	Menggambarkan hubungan antar simbol yang menyatakan suatu jalannya proses dalam sistem.
3		<i>Process</i>	Menggambarkan deskripsi dari proses yang dijalankan.
4		<i>Data</i>	Menggambarkan proses yang dilakukan yang memiliki data masukan (<i>input</i>) atau keluaran (<i>output</i>)
5		<i>Decision</i>	Menggambarkan proses yang memiliki kondisi dimana harus memilih lebih dari satu proses.