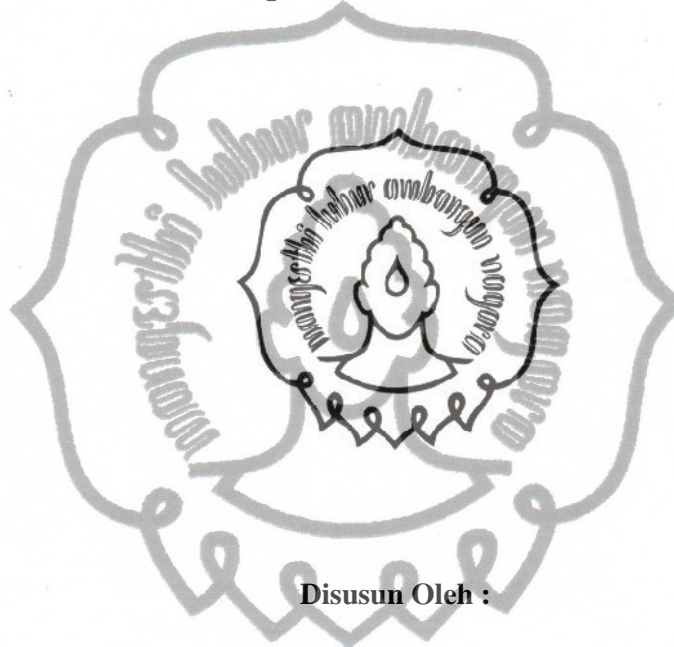


**ANALISIS SENTIMEN MASYARAKAT TERHADAP OPERATOR
SELULER MENGGUNAKAN METODE *SUPPORT VECTOR MACHINE***

SKRIPSI

**Diajukan untuk memenuhi persyaratan mendapatkan gelar Strata Satu
Program Studi Informatika**



Disusun Oleh :

NOVERA ANGGRAENI

M0514038

PROGRAM STUDI INFORMATIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS SEBELAS MARET

2021

**HALAMAN PERSETUJUAN
SKRIPSI**

**ANALISIS SENTIMEN MASYARAKAT TERHADAP OPERATOR SELULER
MENGUNAKAN METODE SUPPORT VECTOR MACHINE**

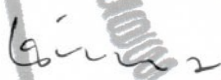
Disusun Oleh:

**NOVERA ANGGRAENI
M0514038**

Skripsi ini telah disetujui untuk dipertahankan di hadapan dewan penguji
pada tanggal, 20 April 2021

Pembimbing I

Pembimbing II



Dr. WIRANTO, M.Kom., M.Cs.
NIP. 196612301993021001

HERI PRASETYO, S.Kom, MSc.Eng,
PhD
NIP. 1983030220161001

HALAMAN PENGESAHAN
SKRIPSI

Analisis Sentimen Masyarakat terhadap Operator Seluler Menggunakan Metode
Support Vector Machine

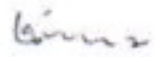
Disusun Oleh:

NOVERA ANGGRAENI
M0514038

Telah dipertahankan di hadapan dewan penguji
pada tanggal 27 April 2021

Susunan Dewan penguji

1. Dr. WIRANTO, M.Kom., M.Cs.
NIP. 196612301993021001

()

2. HERI PRASETYO, S.Kom, MSc.Eng, PhD
NIP. 1983030220161001

()

3. Dr. WIHARTO, S.T., M.Kom.
NIP. 197502102008011005

()

4. Dr. UMI SALAMAH, S.Si., M.Kom.
NIP. 197002171997022001

()

Disahkan oleh

Kepala Program Studi Informatika

commit to user



Dr. WIHARTO, S.T., M.Kom.
NIP. 197502102008011005

ANALISIS SENTIMEN MASYARAKAT TERHADAP OPERATOR SELULER MENGGUNAKAN METODE SUPPORT VECTOR MACHINE

NOVERA ANGGRAENI

Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan
Alam, Universitas Sebelas Maret

ABSTRAK

Saat ini masyarakat cenderung menggunakan media sosial untuk mengungkapkan segala sesuatu termasuk memberikan opini terhadap suatu produk, salah satunya Twitter. *Provider* telekomunikasi menjadi salah satu yang juga membuat akun resmi untuk berinteraksi serta mengetahui tanggapan mengenai produk yang disediakan. Penelitian ini bertujuan untuk melakukan klasifikasi *tweet* opini masyarakat terhadap 3 *provider* telekomunikasi besar di Indonesia, yaitu Indosat, Telkomsel, dan XL dengan menggunakan metode *Support Vector Machine* atau SVM pada media sosial Twitter. Metode *Support Vector Machine* yang digunakan adalah SVM *Multiclass* dengan teknik *One-vs-One* kernel RBF. Melalui pengujian dalam penelitian ini dapat diketahui bahwa tahapan *preprocessing* berpengaruh pada proses klasifikasi, sehingga klasifikasi dilakukan dengan melalui semua tahap *preprocessing*. Didapatkan hasil akurasi sebesar 0.92832 untuk *provider* Indosat, 0.95770 untuk *provider* Telkomsel, dan 0.94444 untuk *provider* XL. Dan dari hasil evaluasi dapat dianalisa bahwa sistem dapat mengklasifikasi kelas negatif dengan baik untuk semua *provider*. Dan sebaliknya untuk kelas yang paling susah diprediksi oleh sistem adalah kelas netral. Dari data hasil klasifikasi, diperoleh beberapa kata yang paling berkontribusi di setiap kelas, sehingga dapat diketahui apa yang banyak dikeluhkan dan banyak membuat masyarakat puas terhadap ketiga *provider*.

Kata kunci: Analisis Sentimen, *Preprocessing*, *Provider* Telekomunikasi, *Support Vector Machine*

SENTIMENT ANALYSIS OF TELECOMMUNICATION SERVICE PROVIDERS WITH SUPPORT VECTOR MACHINE ALGORITHM

NOVERA ANGGRAENI

Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan
Alam, Universitas Sebelas Maret

ABSTRACT

Nowadays, people tend to use social media for express everything, including giving opinions on some products, one of them is Twitter. Telecommunication providers are the ones that also create an official accounts to interact dan get feedback about the products. This research aims to classify public opinion tweets from 3 major telecommunication providers in Indonesia, namely Indosat, Telkomsel, and XL using Support Vector Machine or SVM method on Twitter. The method used is SVM Multiclass with One-vs-One using kernel RBF technique. Through testing in this study, it can be seen that the preprocessing steps has an effect on the classification process, so that classification is carried out through all preprocessing steps. The accuracy results are 0.92832 for Indosat provider, 0.95770 for Telkomsel provider, and 0.94444 for XL provider. From evaluation results, can be analyzed that system can classify negative class well, and for the class the system most difficult to predict is neutral. From classification results data, it is obtained several words that most contribute in each class, so that it can be seen what people complain about and make public satisfied with the providers.

Keywords: *Sentiment Analysis, Preprocessing, Support Vector Machine, Telecommuniacation Providers*

MOTTO

A dream doesn't become reality through magic; it takes sweat, determination, and hard work. (Colin Powel)

Mimpi tidak terwujud nyata melalui ilmu sihir; dibutuhkan keringat, tekad, dan kerja keras. (Colin Powell)



PERSEMBAHAN

Skripsi ini saya persembahkan untuk :

Kedua orang tua yang menjadi alasan utama saya untuk terus melangkah maju.



KATA PENGANTAR

Puji syukur penulis panjatkan kehadirat Allah SWT atas segala rahmat serta hidayah-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul “Analisis Sentimen Masyarakat terhadap Operator Seluler Menggunakan Metode Support Vector Machine”.

Penulis mengucapkan terima kasih kepada semua pihak yang telah membantu dalam penyusunan skripsi ini, terlebih penulis ucapkan terima kasih kepada :

1. Kedua orang tua yang selalu memberikan restu serta doa sehingga penulis dapat menyelesaikan skripsi ini.
2. Bapak Dr. Wiharto, S.T., M.Kom., selaku Kepala Program Studi Informatika FMIPA UNS yang telah memberikan dukungan selama proses penyusunan skripsi.
3. Bapak Dr. Wiranto, M.Kom., M.Cs., dan Bapak Heri Prasetyo, S.Kom., M.Sc.Eng., Ph.D., selaku dosen pembimbing I dan pembimbing II atas ilmu dan masukan yang telah diberikan, serta kesediaan untuk membimbing penulis selama penyusunan skripsi.
4. Teman-teman S1 Informatika yang telah membantu dan memberikan dukungan kepada penulis selama penyusunan skripsi.

Penyusunan skripsi ini tentu masih memiliki beberapa kekurangan, sehingga kritik dan saran dari pembaca masih sangat diperlukan. Semoga skripsi ini bermanfaat bagi pembaca.

Surakarta,

Penulis

DAFTAR ISI

HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN	iii
ABSTRAK	iv
ABSTRACT	v
MOTTO	vi
PERSEMBAHAN	vii
KATA PENGANTAR	viii
DAFTAR ISI	ix
DAFTAR GAMBAR	xi
DAFTAR TABEL	xii
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Perumusan Masalah	2
1.3 Pembatasan Masalah	3
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.6 Sistematika Penulisan	3
BAB 2 TINJAUAN PUSTAKA	5
2.1 Dasar Teori	5
2.1.1 Text Mining	5
2.1.2 Analisis Sentimen	8
2.1.3 Twitter	9
2.1.4 Support Vector Machine (SVM)	9
2.1.5 Evaluasi	14

2.2	Penelitian Terkait.....	16
BAB 3 METODOLOGI PENELITIAN		23
3.1	Pengumpulan Data.....	23
3.2	Pengolahan Data.....	24
3.2.1	Preprosesing Data.....	24
3.2.2	Pembobotan TF-IDF.....	26
3.3	Klasifikasi SVM	26
3.4	Evaluasi.....	27
BAB 4 HASIL DAN PEMBAHASAN		28
4.1	Pengumpulan Data.....	28
4.2	Pengolahan Data.....	29
4.3	Klasifikasi SVM	31
4.4	Evaluasi.....	41
4.5	Wordcloud.....	48
BAB 5 PENUTUP		50
5.1	Kesimpulan	50
5.2	Saran	50

DAFTAR GAMBAR

Gambar 2.1 Tahapan <i>Text Processing</i>	5
Gambar 2.2 Contoh Penggunaan Metode <i>One-vs-One</i> untuk Klasifikasi Data Baru	10
Gambar 3.1 Metodologi Penelitian.....	23
Gambar 3.2 Contoh Penerapan <i>Case Folding</i>	25
Gambar 3.3 Contoh Penerapan <i>Tokenizing</i>	25
Gambar 3.4 Contoh Penerapan <i>Filtering</i>	25
Gambar 3.5 Contoh Penerapan <i>Stemming</i>	26
Gambar 4.1 Kata yang paling berkontribusi dikelas negatif untuk <i>provider</i> Indosat	36
Gambar 4.2 Kata yang paling berkontribusi dikelas negatif untuk <i>provider</i> Telkomsel	36
Gambar 4.3 Kata yang paling berkontribusi dikelas netral untuk <i>provider</i> XL....	37
Gambar 4.4 Kata yang paling berkontribusi dikelas netral untuk <i>provider</i> Indosat	38
Gambar 4.5 Kata yang paling berkontribusi dikelas netral untuk <i>provider</i> Telkomsel	38
Gambar 4.6 Kata yang paling berkontribusi dikelas netral untuk <i>provider</i> XL....	38
Gambar 4.7 Kata yang paling berkontribusi dikelas positif untuk <i>provider</i> Indosat	39
Gambar 4.8 Kata yang paling berkontribusi dikelas positif untuk <i>provider</i> Telkomsel	40
Gambar 4.9 Kata yang paling berkontribusi dikelas positif untuk <i>provider</i> XL....	40
Gambar 4.10 <i>Wordcloud</i> <i>provider</i> Indosat.....	48
Gambar 4.11 <i>Wordcloud</i> <i>provider</i> Telkomsel.....	49
Gambar 4.12 <i>Wordcloud</i> <i>provider</i> XL.....	49

DAFTAR TABEL

Tabel 2.1 Contoh Metode <i>One-vs-All</i> dengan 3 Kelas	10
Tabel 2.2 Contoh Metode <i>One-vs-One</i> dengan 3 Kelas	10
Tabel 2.3 Jenis Fungsi Kernel yang umum dipakai dalam SVM	13
Tabel 2.4 Confusion Matrix Klasifikasi 3 Kelas	15
Tabel 2.5 Penelitian Terkait	20
Tabel 3.1 Daftar Akun Resmi Operator untuk Pengumpulan Data	24
Tabel 3.2 Karakteristik klasifikasi manual	24
Tabel 4.1 Dataset <i>Tweet</i> Provider	28
Tabel 4.2 Perbandingan <i>Tweet</i> Sebelum dan Sesudah <i>Case Folding</i>	29
Tabel 4.3 Perbandingan <i>Tweet</i> Sebelum dan Sesudah Pembersihan Atribut <i>Tweet</i>	29
Tabel 4.4 Perbandingan <i>Tweet</i> Sebelum dan Sesudah Penghilangan Karakter yang Tidak Diperlukan	30
Tabel 4.5 Perbandingan <i>Tweet</i> Sebelum dan Sesudah Penghapusan <i>Stopword</i>	30
Tabel 4.6 Perbandingan <i>Tweet</i> Sebelum dan Sesudah <i>Stemming</i>	31
Tabel 4.7 Hasil klasifikasi dengan data hasil <i>preprocessing</i> tanpa <i>filtering</i> dan <i>stemming</i>	32
Tabel 4.8 Hasil klasifikasi dengan melakukan tahap <i>preprocessing</i> secara keseluruhan	32
Tabel 4.9 Hasil Akurasi Klasifikasi SVM untuk <i>Provider</i> Indosat dengan Data Latih 20% (a), Data Latih 30% (b), dan Data Latih 40% (c)	33
Tabel 4.10 Hasil Akurasi Klasifikasi SVM untuk <i>Provider</i> Telkomsel dengan Data Latih 20% (a), Data Latih 30% (b), dan Data Latih 40% (c)	34
Tabel 4.11 Hasil Akurasi Klasifikasi SVM untuk <i>Provider</i> XL dengan Data Latih 20% (a), Data Latih 30% (b), dan Data Latih 40% (c)	35
Tabel 4.12 <i>Confusion Matrix</i> untuk <i>provider</i> Indosat	41
Tabel 4.13 <i>Confusion Matrix</i> untuk <i>provider</i> Telkomsel	43
Tabel 4.14 <i>Confusion Matrix</i> untuk <i>provider</i> XL	45