

SKRIPSI
SENTIMEN ANALISIS SISTEM ZONASI SEKOLAH PADA MEDIA
SOSIAL TWITTER MENGGUNAKAN METODE *SUPPORT*
VECTOR MACHINE



AKHIRA MAULIDIO FIRDAZA

Nomor Mahasiswa : 175410109

PROGRAM STUDI TEKNIK INFORMATIKA
SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
AKAKOM YOGYAKARTA

2021

SKRIPSI

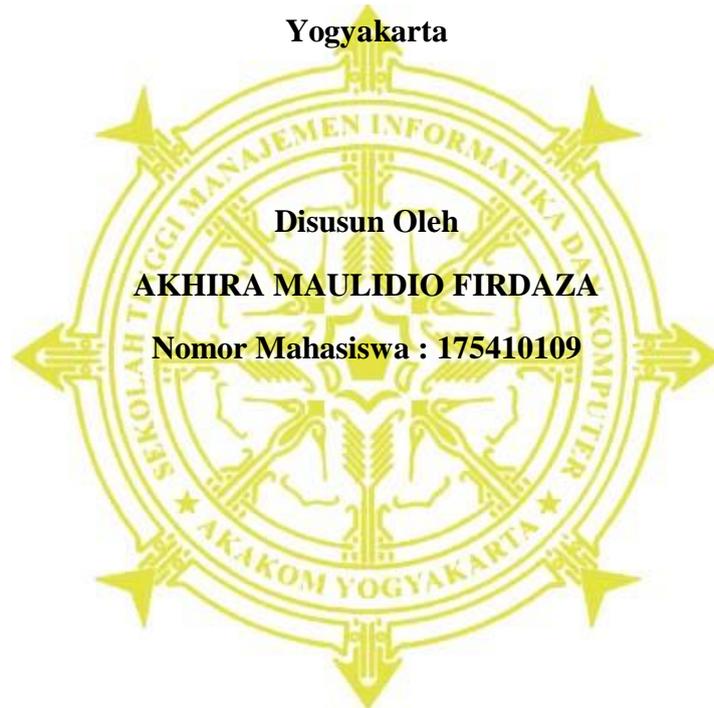
**SENTIMEN ANALISIS SISTEM ZONASI SEKOLAH PADA MEDIA
SOSIAL TWITTER MENGGUNAKAN METODE *SUPPORT*
*VECTOR MACHINE***

**Diajukan sebagai salah satu syarat untuk menyelesaikan studi jenjang strata satu
(S1) Program Studi Teknik Informasi**

Sekolah Tinggi Manajemen Informatika dan Komputer

Akakom

Yogyakarta



PROGRAM STUDI TEKNIK INFORMATIKA

**SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
AKAKOM YOGYAKARTA**

2021

HALAMAN PERSETUJUAN

Judul : SENTIMEN ANALISIS SISTEM ZONASI SEKOLAH PADA MEDIA
SOSIAL TWITTER MENGGUNAKAN METODE *SUPPORT VECTOR
MACHINE*

Nama : Akhira Maulidio Firdaza

NIM : 175410109

Jurusan : Informatika

Semester : Tujuh (7)

Telah memenuhi syarat dan disetujui untuk diselenggarakan di
Hadapan dosen penguji seminar tugas akhir

Yogyakarta, _____ 2021

Dosen pembimbing,



Ariesta Damayanti, S.Kom., M.Cs.

HALAMAN PENGESAHAN

SKRIPSI

**SENTIMEN ANALISIS SISTEM ZONASI SEKOLAH PADA MEDIA SOSIAL
TWITTER MENGGUNAKAN METODE *SUPPORT VECTOR MACHINE***

Telah dipertahankan dan diujikan di depan Dewan Penguji Skripsi dan dinyatakan diterima sebagai syarat guna memperoleh Gelar Sarjana Komputer Sekolah Tinggi

Manajemen Informatika dan Komputer

YOGYAKARTA

Yogyakarta, 2021

Mengesahkan

Dewan Penguji

1. Dini Fakta Sari, S.T., M.T.
2. Ariesta Damayanti, S.Kom., M.Cs.

Tanda Tangan



Mengetahui,

Ketua Program Studi Teknik Informatika

18 FEB 2021



Dini Fakta Sari, S.T., M.T.

HALAMAN PERSEMBAHAN

Allhamdulillahirobbil'alamin,

Sujud Syukur ku panjatkan kepada Allah SWT, yang telah memberikan nikmat serta ridhonya sehingga saya dapat menyelesaikan skripsi ini. Semoga dengan selesainya skripsi ini menjadi langkah awal untuk tercapainya impian dan cita – cita saya.

Dengan kerendahan hati skripsi ini kupersembahkan Kepada :

1. Papa dan Mama tersayang.

Sebagai Orang Tua yang telah mendidik saya dengan penuh kasih sayang sejak dilahirkan sampai dengan saat ini. Terimakasih karena tidak pernah lelah membantu, mendoakan dan selalu *support* Rara untuk selalu mewujudkan cita-cita saya.

2. Teman-teman

Sebagai teman-teman yang selalu mendorong, memberi semangat, memotivasi, tempat curhat, berbagi info terkini. Rara mengucapkan terimakasih telah memberi warna dalam kehidupan perkuliahan saya. ABAB CERIA (Indah, Ramzy, Mami Khansa Ulil, Bondan, Akbar, Panji, Ulum, Akim, Kibay) ,Team SUJUD Company , Bli Kadek, dan Hilda teman seperjuangan .

3. Kepada Sodikin aka Manan yang selalu memberikan *support* dan siap membantu setiap saat ketika dibutuhkan.

4. Badan Eksekutif Mahasiswa (BEM)

Organisasi kampus yang dapat menjadi keluarga, rumah, dan tempat belajar kedua saya selama di Yogyakarta. Memberikan saya banyak sekali pengalaman yang tak terlupakan selama saya menjadi anggota.

Memperdalam *soft skills*, keberanian, kepribadian, dan rasa tanggung jawab saya sehingga saya dapat mengimplementasikan pengalaman tersebut dimasa depan.

5. Tidak lupa saya ucapkan terima kasih banyak kepada Ibu Ariesta Damayanti S.Kom., M.Cs. sebagai dosen pembimbing yang telah bersedia membimbing saya, dan juga semua dosen STMIK AKAKOM telah menularkan ilmu yang kalian miliki. Semoga ilmu yang kalian berikan dapat bermanfaat.

MOTTO

“lebih baik merangkak sedikit demi sedikit daripada berhenti berjalan”

INTISARI

Adanya sistem zonasi dalam sistem Penerimaan Peserta Didik Baru (PPDB) di Indonesia menyebabkan banyak masyarakat yang menuangkan opini mereka tentang sistem zonasi yang ada di Indonesia melalui media sosial. Salah satu media sosial yang populer di masyarakat adalah Twitter. Media sosial ini mampu menghubungkan banyak orang dengan berbagai topik dari seluruh dunia. Dengan menggunakan Twitter masyarakat dapat memberikan berbagai macam pendapat mereka tentang apapun dalam bentuk komentar.

Dalam penelitian ini, diklasifikasikan apakah suatu *tweet* termasuk dalam kelas positif, negatif atau netral dalam kaitan penerapan sistem zonasi di Indonesia. Oleh karena itu, *tweet* yang telah dikumpulkan dan diberi label perlu dipraproses, meliputi *case folding*, *cleaning*, *tokenisasi*, normalisasi kata, *stopword removal* dan *stemming*. Selanjutnya, dilakukan ekstraksi fitur berdasarkan nilai *document frequency* dan *threshold minimum* bagi kata untuk dipilih sebagai fitur.

Setiap ditentukan nilai fitur dengan term *frequency-inverse document frequency*. Setelah dilakukan percobaan metode *Support Vector Machine*(SVM) menggunakan kernel linear dan percobaan tuning parameter dengan menggunakan *kernel RBF*, diketahui hasil akurasi dengan menggunakan kernel linear adalah 75% . Parameter terbaik adalah pasangan parameter $C=1$ dan $\gamma=1$ menggunakan *kernel RBF* dengan rata-rata akurasi sebesar 81% .

Kata kunci – *kernel RBF* , *klasifikasi* , *Sistem zonasi*, *SVM* , *Tweet*,

KATA PENGANTAR

Segala puji dan syukur kehadirat Allah SWT atas berkah, rahamat dan hidayah-Nya yang senantiasa dilimpahkan kepada penulis, sehingga penulis dapat menyelesaikan penyusunan skripsi dengan judul “Sentimen Analisis Sistem Zonasi Sekolah Pada Media Sosial Twitter Menggunakan Metode *Support Vector Machine*”. Sebagai salah satu syarat menyelesaikan studi jenjang Strata Satu (S1) Program studi Teknik Informatika di Sekolah Tinggi Manajemen Informatika dan Komputer AKAKOM Yogyakarta.

Dalam penyusunan dan penulisan ini tidak terlepas dari bantuan, bimbingan serta dukungan dari berbagai pihak. Oleh karena itu, tanpa mengurangi rasa hormat dengan segala kerendahan hati penulis ingin menyampaikan rasa terimakasih kepada:

1. Bapak Ir. Totok Suprawoto, M.M., M.T. selaku Ketua STMIK AKAKOM Yogyakarta.
2. Ibu Dini Fakta Sari, S.T., M.T. selaku Ketua Program Studi Teknik Informatika Sekolah Tinggi Manajemen Informatika dan Komputer Akakom Yogyakarta.
3. Ibu Ariesta Damayanti, S.Kom., M.Cs selaku dosen pembimbing yang telah memberikan bimbingan, arahan, saran serta waktunya selama penelitian dan penulisan skripsi ini.
4. Semua pihak yang telah membantu, baik secara langsung maupun tidak langsung yang tidak dapat penulis sebutkan satu persatu.

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERSEMBAHAN	iv
MOTTO	vi
INTISARI	vii
KATA PENGANTAR	viii
DAFTAR ISI	ix
DAFTAR GAMBAR	xi
DAFTAR TABEL	xii
BAB I PENDAHULUAN	1
1.1. Latar belakang	1
1.2. Rumusan Masalah	4
1.3. Ruang Lingkup	5
1.4. Tujuan Penelitian.....	6
1.5. Manfaat Penelitian.....	6
BAB II TINJAUAN PUSTAKA DAN DASAR TEORI	7
2.1. Tinjauan Pustaka	7
2.2. Dasar Teori	10
2.2.1. Twitter	10
2.2.2. Analisis Sentimen	10
2.2.3. <i>Preprocessing</i>	11
2.2.4. <i>Term Frequency-Inverse Document Frequency</i>	12
2.2.5. <i>Support Vector Machine</i>	14
2.2.6. <i>Pandas Library</i>	15
2.2.7. <i>Scikit-learn</i>	16
BAB III METODE PENELITIAN	16
3.1. Bahan/Data	16
3.1.1. Kebutuhan Input.....	16

3.1.2.	Kebutuhan Output	16
3.2.	Peralatan.....	17
3.2.1.	Perangkat Lunak.....	17
3.2.2.	Perangkat keras.....	17
3.3.	Prosedur dan Pengumpulan Data	17
3.3.1.	Metode Pengumpulan Data	17
3.3.2.	Prosedur Pengumpulan Data	19
3.4.	Analisis dan Perancangan Sistem.....	19
3.4.1.	Proses pemasukan data.....	20
3.4.2.	Pelabelan Data Manual.....	21
3.4.3.	Tahap <i>Preprocessing</i>	21
3.4.4.	Pembobotan TF-IDF.....	23
3.4.5.	Proses SVM Training	26
3.4.6.	Proses SVM Testing	27
3.4.7.	Proses SVM Training dengan <i>Kernel Radial Basis Function</i>	28
3.4.8.	Rancangan Antar Muka	29
	Usecase Diagram	29
	Tampilan Home	30
	Tampilan Load Data	30
	Tampilan Data <i>Preprocessing</i>	31
	Tampilan Hasil Pembobotan	31
	BAB IV IMPLEMENTASI DAN PEMBAHASAN	33
4.1.	Implementasi dan Uji Coba Sistem.....	33
4.2.	Pembahasan	39
	BAB V PENUTUP	60
5.1.	Kesimpulan	60
5.2.	Saran	61
	DAFTAR PUSTAKA	62
	LAMPIRAN	

DAFTAR GAMBAR

Gambar 2. 1 Ilustrasi metode <i>Support Vector Machine</i>	15
Gambar 3. 1 Diagram Alur sentimen analisis dengan Support Vector Machine..	20
Gambar 3. 2 Diagram Alur sentimen analisis dengan Support Vector Machine menggunakan kernel Radial Basis Function	21
Gambar 3. 3 Ilustrasi metode Train Test Split	27
Gambar 3. 4 Use Case Diagram	30
Gambar 3. 5 Tampilan Home.....	31
Gambar 3. 6 Halaman Load Data	31
Gambar 3. 7 Tampilan Data Preprocessing	32
Gambar 3. 8 Tampilan Hasil Pembobotan TF-IDF	32
Gambar 3. 9 Tampilan Output berupa diagram pie.....	33
Gambar 4. 1 potongan program crawling menggunakan tool twint	34
Gambar 4. 2 Potongan Kode Program Case Folding.....	35
Gambar 4. 3 Potongan Kode Program Tokenizing	38
Gambar 4. 4 Potongan Kode Program <i>Stopword Removal</i>	39
Gambar 4. 5 Potongan Kode Program Normalisasi	40
Gambar 4. 6 Potongan Kode Program <i>Stemming</i>	42
Gambar 4. 7 Potongan Kode Program Case Folding.....	42
Gambar 4. 8 Potongan Kode Program Proses SVM Training	43
Gambar 4. 9 Potongan Kode Program SVM Testing	43
Gambar 4. 10 Data Hasil Crawling dalam Format .csv	44
Gambar 4. 11 <i>Mapping Term</i>	50

DAFTAR TABEL

Tabel 2. 1 Perbandingan Penelitian	9
Tabel 3. 1 <i>Preprocessing data</i>	23
Tabel 3. 2 Term dari Data Tweet.....	24
Tabel 3. 3 Hasil perhitungan TF.....	25
Tabel 3. 4 Hasil Perhitungan TF-IDF	26
Tabel 3. 5 Tabel Confussion Matrix.....	28
Tabel 4. 1. Pelabelan biner	44
Tabel 4. 2 Hasil Case folding.....	45
Tabel 4. 3. Hasil Tokenizing.....	46
Tabel 4. 4 Hasil <i>Stop Word</i>	48
Tabel 4. 5. Hasil Normalisasi.....	50
Tabel 4. 6. Hasil <i>Stemming</i>	51
Tabel 4. 7. Potongan Hasil TF-IDF	54
Tabel 4. 8. Potongan Hasil TF-IDF	55
Tabel 4. 9. Tabel perbandingan actual class & predicted class confusion matrix	57
Tabel 4. 10. Diagram Pie presentase sentimen dengan kernel RBF	58
Tabel 4. 11. Hasil parameter terbaik pasangan c dan γ	59
Tabel 4. 12. Tabel perbandingan actual class & predicted class confusion matrix	60
Tabel 4. 13. Confusion Matrix	64
Tabel 4. 14. Diagram Pie presentase sentimen dengan kernel RBF Hasil.....	61