

**ANALISIS SENTIMEN OPINI MASYARAKAT INDONESIA
TERHADAP PANDEMI VIRUS CORONA (*COVID-19*) DI
MEDIA SOSIAL *TWITTER* MENGGUNAKAN METODE
*SUPPORT VECTOR MACHINE***

SKRIPSI

RAFIZAH AINI

0701162014



**PROGRAM STUDI ILMU KOMPUTER
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SUMATERA UTARA
MEDAN**

2023

**ANALISIS SENTIMEN OPINI MASYARAKAT INDONESIA
TERHADAP PANDEMI VIRUS CORONA (COVID-19) DI
MEDIA SOSIAL *TWITTER* MENGGUNAKAN METODE
*SUPPORT VECTOR MACHINE***

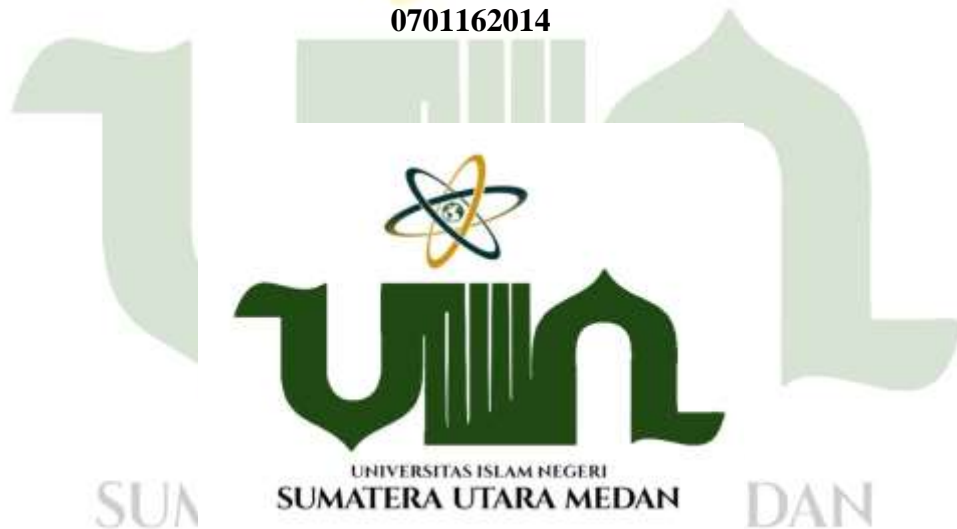
SKRIPSI

Diajukan untuk Memenuhi Syarat Mencapai Gelar Sarjana Komputer



RAFIZAH AINI

0701162014



**PROGRAM STUDI ILMU KOMPUTER
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SUMATERA UTARA
MEDAN**

2023

PERSETUJUAN SKRIPSI

Hal : Surat Persetujuan Skripsi

Lamp : -

Kepada Yth.,

Dekan Fakultas Sains dan Teknologi

Universitas Islam Negeri Sumatera Utara Medan

Assalamu'alaikum Wr. Wb

Setelah membaca, meneliti, memberikan petunjuk, dan mengoreksi serta mengadakan perbaikan, maka kami selaku pembimbing berpendapat bahwa skripsi saudara

Nama : Rafizah Aini

Nomor Induk Mahasiswa : 0701162014

Program Studi : Ilmu Komputer

Judul : Analisis Sentimen Opini Masyarakat Indonesia Terhadap Pandemi Virus Corona (*Covid-19*) Di Media Sosial *Twitter* Menggunakan Metode *Support Vector Machine*

dapat disetujui untuk segera dimunaqasyahkan. Atas perhatiannya kami ucapkan terimakasih.

Medan, 16 Januari 2023 M

23 Jumadil Akhir 1444 H

Komisi Pembimbing

Pembimbing I,



Dr. Mhd. Furqan, S.Si., M.Comp.Sc.
NIP. 198008062006041003

Pembimbing II,



Muhammad Ikhsan, S.T., M.Kom
NIP. 198304152011011008

SURAT PERNYATAAN KEASLIAN SKRIPSI

Saya yang bertanda tangan di bawah ini,

Nama : Rafizah Aini

Nomor Induk Mahasiswa : 0701162014

Program Studi : Ilmu Komputer

Judul : Analisis Sentimen Opini Masyarakat Indonesia Terhadap Pandemi Virus Corona (*Covid-19*) Di Media Sosial *Twitter* Menggunakan Metode *Support Vector Machine*

menyatakan bahwa skripsi ini adalah hasil karya saya sendiri, kecuali beberapa kutipan dan ringkasan yang masing-masing disebutkan sumbernya. Apabila di kemudian hari ditemukan plagiat dalam skripsi ini maka saya bersedia menerima sanksi pencabutan gelar akademik yang saya peroleh dan sanksi lainnya sesuai dengan peraturan yang berlaku.

Medan, 16 Januari 2023



Rafizah Aini

NIM. 0701162014

PENGESAHAN SKRIPSI

Nomor: B.048/ST/ST.V.2/PP.01.1/02/2023

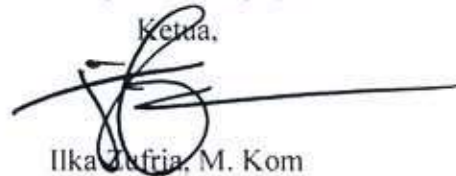
Judul : Analisis Sentimen Opini Masyarakat Indonesia Terhadap Pandemi Virus Corona (*Covid-19*) Di Media Sosial *Twitter* Menggunakan Metode *Support Vector Machine*
Nama : Rafizah Aini
Nomor Induk Mahasiswa : 0701162014
Program Studi : Ilmu Komputer
Fakultas : Sains dan Teknologi

Telah dipertahankan di hadapan Dewan Penguji Skripsi Program Studi Ilmu Komputer Fakultas Sains dan Teknologi UIN Sumatera Utara Medan dan dinyatakan **LULUS**.

Pada hari/tanggal : Rabu/ 22 Februari 2023
Tempat : Ruang Meeting Fakultas Sains dan Teknologi UIN Sumatera Utara Medan – Kampus IV Tuntungan

Tim Ujian Munaqasyah,

Ketua,



Ilka Zulfria, M. Kom
NIP. 198506042015031006

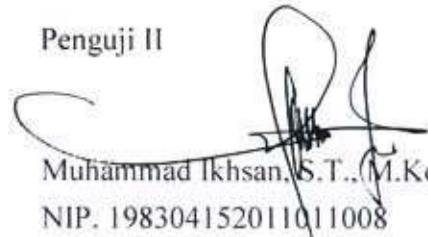
Dewan Penguji,

Penguji I,



Dr. Mhd. Furqan, S.Si, M.Comp.Sc.
NIP. 198008062006041003

Penguji II



Muhammad Ikhsan, S.T., M.Kom
NIP. 198304152011011008

Penguji III,



Ilka Zulfria, M. Kom
NIP. 198506042015031006

Penguji IV



Rakhmat Kurniawan R, S.T., M.Kom
NIP. 198503162015031003

Mengesahkan,
Dekan Fakultas Sains dan Teknologi
UIN Sumatera Utara Medan,



Prof. Dr. H. H. Syahnan, M.A
NIP. 196609051991031002

ABSTRAK

Pandemi virus corona adalah peristiwa menyebarnya penyakit koronavirus 2019 atau *coronavirus disease 2019* di seluruh dunia. Penyakit ini disebabkan oleh koronavirus jenis baru yang diberi nama SARS-CoV-2. Pandemi yang terjadi di seluruh dunia menjadi topik pembicaraan, termasuk di media sosial. Salah satu media sosial yang sering digunakan masyarakat adalah *twitter*. Di media sosial *twitter* pandemi virus corona selalu menjadi topik pembicaraan yang sering di bahas hingga menimbulkan kontroversi. Kontroversi terjadi dikarenakan setiap harinya opini-opini di media sosial *twitter* terkait pandemi virus corona selalu bertambah sehingga, ketika masyarakat membaca berita di media sosial mengenai pandemi tersebut menimbulkan kekhawatiran dikarenakan opini masyarakat yang berbeda-beda. Dari permasalahan ini penulis akan membuat sebuah sistem yang menganalisis opini dari media sosial *twitter* untuk mendapatkan sentimen opini apa yang terjadi pada masyarakat mengenai permasalahan pandemi virus corona tersebut. Penelitian ini menggunakan metode *support vector machine*. *Support Vector Machine* merupakan metode yang cepat dan efektif untuk klasifikasi teks. Hasil penelitian ini akan melakukan klasifikasi kalimat positif, negatif dan netral. Akurasi yang didapatkan dari algoritma *Support Vector Machine* 98%. Pengujian dilakukan dengan perhitungan *precision*, *recall*, *F-measure*.

Kata kunci: Twitter, opini, algoritma, klasifikasi

ABSTRACT

The coronavirus pandemic is an event that causes the spread of the coronavirus disease 2019 or coronavirus disease 2019 around the world. This disease is caused by a new type of coronavirus called SARS-CoV-2. The pandemic that has occurred around the world has become a topic of conversation, including on social media. One of the social media that is often used by the public is Twitter. On social media, the corona virus pandemic has always been a topic of conversation that is often discussed, causing controversy. Controversy occurs because every day the opinions on social media Twitter regarding the corona virus pandemic are always increasing so that, when people read news on social media about the pandemic, it raises concerns because people's opinions are different. From these problems the author will create a system that analyzes opinions from Twitter social media to get opinion sentiment about what is happening in the community regarding the problem of the corona virus pandemic. This research uses the support vector machine method. Support Vector Machine is a fast and effective method for text classification. The results of this study will classify positive, negative and neutral sentences. The accuracy obtained from the Support Vector Machine algorithm is 98%. Testing is done by calculating precision, recall, F-measure.

Keyword: Twitter, opinion, algorithm, classification



UNIVERSITAS ISLAM NEGERI
SUMATERA UTARA MEDAN

KATA PENGANTAR

Puji dan Syukur penulis ucapkan Kepada Allah Subhanahu Wa Ta'ala, yang telah memberikan Nikmat dan Rahmat-Nya sehingga penulis dapat menyelesaikan skripsi ini yang berjudul: “Analisis Sentimen Opini Masyarakat Indonesia Terhadap Pandemi Virus Corona (*Covid-19*) Di Media Sosial *Twitter* Menggunakan Metode *Support Vector Machine*”. Sebagai salah satu persyaratan untuk memperoleh gelar sarjana program studi Ilmu Komputer Universitas Islam Negeri Sumatera Utara. Shalawat serta salam tidak lupa selalu tercurahkan kepada Nabi Muhammad Shallallahu ‘alaihi wasallam sebagai panutan kita di dalam kehidupan.

Penyelesaian skripsi ini tentunya tidak dapat terselesaikan tanpa bantuan, dukungan, dan kerja sama dari berbagai pihak. Oleh karena itu, penulis menyampaikan ucapan terima kasih yang sebesar-besarnya kepada:

1. Prof. Dr. Abu Rokhmat, M.A, selaku PLT Rektor Universitas Islam Negeri Sumatera Utara.
2. Prof. Dr. Mhd. Syahnan, M.A, selaku Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sumatera Utara.
3. Ilka Zufria, M. Kom, selaku Ketua Program Studi Ilmu Komputer Fakultas Sains dan Teknologi Universitas Islam Negeri Sumatera Utara
4. Dr. Mhd. Furqan, S.Si, M.Comp.Sc, selaku Dosen Pembimbing I
5. Muhammad Ikhsan, S.T., M.Kom selaku Dosen Pembimbing Akademik sekaligus Dosen Pembimbing II
6. Teristimewa untuk Kedua Orangtua Penulis, Ayah Rahmadsyah Putra dan Ibu Chadijah, Terimakasih atas segala doa dan dukungan dalam segala bentuk yang selalu menyertai disetiap langkah kehidupan penulis.

7. Abang Rizki Ramadhan, S.H, Terimakasih atas segala bantuan dan dukungannya.
8. Seluruh Keluarga Penulis, Terimakasih atas segala bantuan dan dukungannya.
9. Teman-Teman Ilmu Komputer yang selalu memberi dukungan.
10. Dosen Ilmu Komputer dan Staff kerja Sub-Bagian Akademik dan Sub-Bagian Umum yang berada di Lingkungan Fakultas Sains dan Teknologi.
11. Serta semua pihak yang tidak dapat saya sebutkan satu persatu.

Penulis juga menyadari skripsi ini masih memiliki kekurangan dan belum sempurna. Untuk itu penulis memohon maaf atas kesalahan dan kekurangan yang terdapat di dalamnya. Penulis berharap semoga skripsi ini dapat bermanfaat bagi semua pembacanya.

Medan, 09 Januari 2023

Penulis,

UNIVERSITAS ISLAM NEGERI
SUMATERA UTARA MEDIAN
RAFIZAH AINI
NIM. 0701162014

DAFTAR ISI

ABSTRAK	i
ABSTRACT	ii
KATA PENGANTAR.....	iii
DAFTAR ISI.....	v
DAFTAR GAMBAR.....	viii
DAFTAR TABEL	ix
DAFTAR LAMPIRAN	x
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	3
1.3 Batasan Masalah	4
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	4
BAB II TINJAUAN PUSTAKA.....	5
2.1 Analisis Sentimen	5
2.2 Pandemi Virus Corona.....	6
2.3 Twitter.....	6
2.4 Twitter API (Application Programming Interface)	7
2.5 Text Mining	8
2.6 Text Preprocessing.....	13
2.7 Klasifikasi	16
2.8 Machine Learning	16
2.9 Support Vector Machine (SVM).....	17

2.10 Confusion Matrix	21
2.11 Python	21
2.12 Flowchart	22
2.13 Unified Modelling language (UML)	23
BAB III METODE PENELITIAN	26
3.1 Waktu dan Jadwal Penelitian	26
3.2 Bahan dan Alat Penelitian	26
3.2.1 Perangkat Keras (<i>Hardware</i>)	27
3.2.2 Perangkat Lunak (<i>Software</i>)	27
3.3 Kerangka kerja penelitian	27
3.3.1 Perencanaan	28
3.3.2 Teknik Pengumpulan Data	29
3.4 Analisis Kebutuhan	31
3.5 Perancangan <i>Flowchart</i> sistem	32
BAB IV HASIL DAN PEMBAHASAN	34
4.1 Analisis Sistem	34
4.1.1 Proses crawling data	35
4.1.2 Analisis Data Masukan	36
4.1.3 Preprocessing Teks	36
4.1.4 Case Folding	37
4.1.5 Filtering	37
4.1.6 Tokenization	38
4.1.7 Word Normalization	39
4.1.8 Analisa Algoritma Support Vector Machine	40
4.2 Perancangan Sistem	46
4.3 Hasil dan Implementasi	51
4.3.1 Proses Akuisisi Data	51

4.3.2 Preprocessing	53
4.3.3 Proses Pembersihan Data.....	53
4.3.4 Case Folding	54
4.4 Distribusi Kata.....	55
4.4.1 Word Cloud.....	56
4.5 Support Vector Machine	58
4.6 Confussion Matrix.....	58
BAB V KESIMPULAN DAN SARAN	61
5.1 Kesimpulan.....	61
5.2 Saran	61
DAFTAR PUSTAKA	62
LAMPIRAN-LAMPIRAN	



DAFTAR GAMBAR

Gambar	Judul Gambar	Halaman
2.1	Proses Text Mining	11
2.2	<i>Struktur Support Vector Machine</i>	18
2.3	<i>Hyperplane terbentuk diantara kelas -1 dan +1</i>	19
2.4	<i>Non-linear SVM</i>	20
2.5	Confusion Matrix	21
3.1	Perencanaan.....	28
3.2	<i>Flowchart Sistem Analisis Sentimen</i>	33
4.1	Gambaran Umum Sistem yang Akan Dibangun	34
4.2	Script Python Crawling	36
4.3	Tahapan <i>Preprocessing</i>	36
4.4	Proses <i>Case Folding</i>	37
4.5	Proses <i>Filtering</i>	38
4.6	Proses <i>Word Normalization</i>	39
4.7	Algoritma <i>Support Vector Machine</i>	41
4.8	<i>Usecase Diagram</i> Penelitian.....	47
4.9	<i>Sequence Diagram</i> Input Data.....	50
4.10	<i>Sequence Diagram</i> Training	50
4.11	<i>Sequence Diagram</i> Testing.....	51
4.12	Menampilkan Hasil <i>Crawling</i>	51
4.13	Distribusi Kata.....	55
4.14	<i>Word Cloud</i> Kata <i>Tweet</i> dengan label positif.....	56
4.15	<i>Word Cloud</i> Kata <i>Tweet</i> dengan label negatif.....	57
4.16	<i>Word Cloud</i> Kata <i>Tweet</i> dengan label Netral.....	57
4.17	Hasil Precision, Recall, dan F1-score.....	59
4.18	<i>Confussion Matrix</i>	59

DAFTAR TABEL

Tabel	Judul Tabel	Halaman
2. 1	Jenis kernel SVM yang umum digunakan.....	20
2. 2	Simbol-Simbol Flowchart	22
2. 3	Notasi Utama Use Case Diagram.....	24
3. 1	Jadwal Pelaksanaan Penelitian	26
3. 2	Data <i>Tweet</i> Covid 19	29
3. 3	Pemisahan Kata	42
3. 4	Pembobotan Kata	42
3. 5	Data Uji	46
4. 1	Case Folding.....	37
4. 2	Filtering	38
4. 3	Tokenization.....	39
4. 4	Word Normalization.....	40
4. 5	Skenario preprocessing	47
4. 6	Skenario training	48
4. 7	Skenario testing	48
4. 8	Skenario SVM	49
4. 9	Data <i>Tweet</i> Covid 19	51
4. 10	Pembersihan data.....	53
4. 11	Proses <i>Case Folding</i>	54
4. 12	Data Pengujian dan Data Pelatihan	58

DAFTAR LAMPIRAN

Lampiran	Judul Lampiran
1.	Hasil Uji Sentiment Analisis Terhadap Data Covid-19 Yang Berasal Dari Media Sosial Twitter
2.	Source Code Dari Bahasa Pemrograman Python Dalam Melakukan Sentiment Analisis Terhadap Data Covid-19 Yang Berasal Dari Media Sosial Twitter
3.	Daftar Riwayat Hidup
4.	Kartu Bimbingan Skripsi



UNIVERSITAS ISLAM NEGERI
SUMATERA UTARA MEDAN