

**ANALISIS SENTIMEN ANIMO MASYARAKAT KABUPATEN
KARANGANYAR TERHADAP VAKSINISASI COVID-19
MENGUNAKAN ALGORITMA SVM**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



diajukan oleh

ALFIAN ADY CHANDRA

18.11.2019

Kepada

PROGRAM SARJANA

PROGRAM STUDI INFROMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS AMIKOM YOGYAKARTA

YOGYAKARTA

2023

**ANALISIS SENTIMEN ANIMO MASYARAKAT KABUPATEN
KARANGANYAR TERHADAP VAKSINISASI COVID-19
MENGUNAKAN ALGORITMA SVM**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana

Program Studi (*Tuliskan Prodi Anda*)



diajukan oleh

ALFIAN ADY CHANDRA

18.11.2019

Kepada

PROGRAM SARJANA

PROGRAM STUDI INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS AMIKOM YOGYAKARTA

YOGYAKARTA

2023

HALAMAN PERSETUJUAN

SKRIPSI

**ANALISIS SENTIMEN ANIMO MASYARAKAT KABUPATEN
KARANGANYAR TERHADAP VAKSINISASI COVID-19 MENGGUNAKAN
ALGORITMA SVM**

yang disusun dan diajukan oleh

Alfian Ady Chandra

18.11.2019

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 15 Agustus 2023

Kusnawi, S.Kom, M.Eng,

ij



Kusnawi, S.Kom, M.Eng

NIK. 190302112

HALAMAN PENGESAHAN

SKRIPSI

**ANALISIS SENTIMEN ANIMO MASYARAKAT KABUPATEN
KARANGANYAR TERHADAP VAKSINISASI COVID-19 MENGGUNAKAN
ALGORITMA SVM**

yang disusun dan diajukan oleh

Alfian Ady Chandra

18.11.2019

Telah dipertahankan di depan Dewan Penguji
pada tanggal 15 Agustus 2023

Susunan Dewan Penguji

Nama Penguji

Sri Ngudi Wahvuni, S.T.
M.Kom.
NIK. 190302060

Theopilus Bayu Sasongko.
S.Kom. M.Eng.
NIK. 190302375

Kusnawi, S.Kom. M.Eng.
NIK. 190302112

Tanda Tangan



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 15 Agustus 2023

DEKAN FAKULTAS ILMU KOMPUTER



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

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Alfian Ady Chandra
NIM : 18.11.2019

Menyatakan bahwa Skripsi dengan judul berikut:

Analisis Sentimen Animo Masyarakat Kabupaten Karanganyar Terhadap Vaksinisasi Covid-19 Menggunakan Metode SVM

Dosen Pembimbing : Kusnawi, S.Kom, M.Eng

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, 15 Agustus 2023

Yang Menyatakan,



Alfian Ady Chandra

MOTTO

“Unbothered alien who lost in slum”

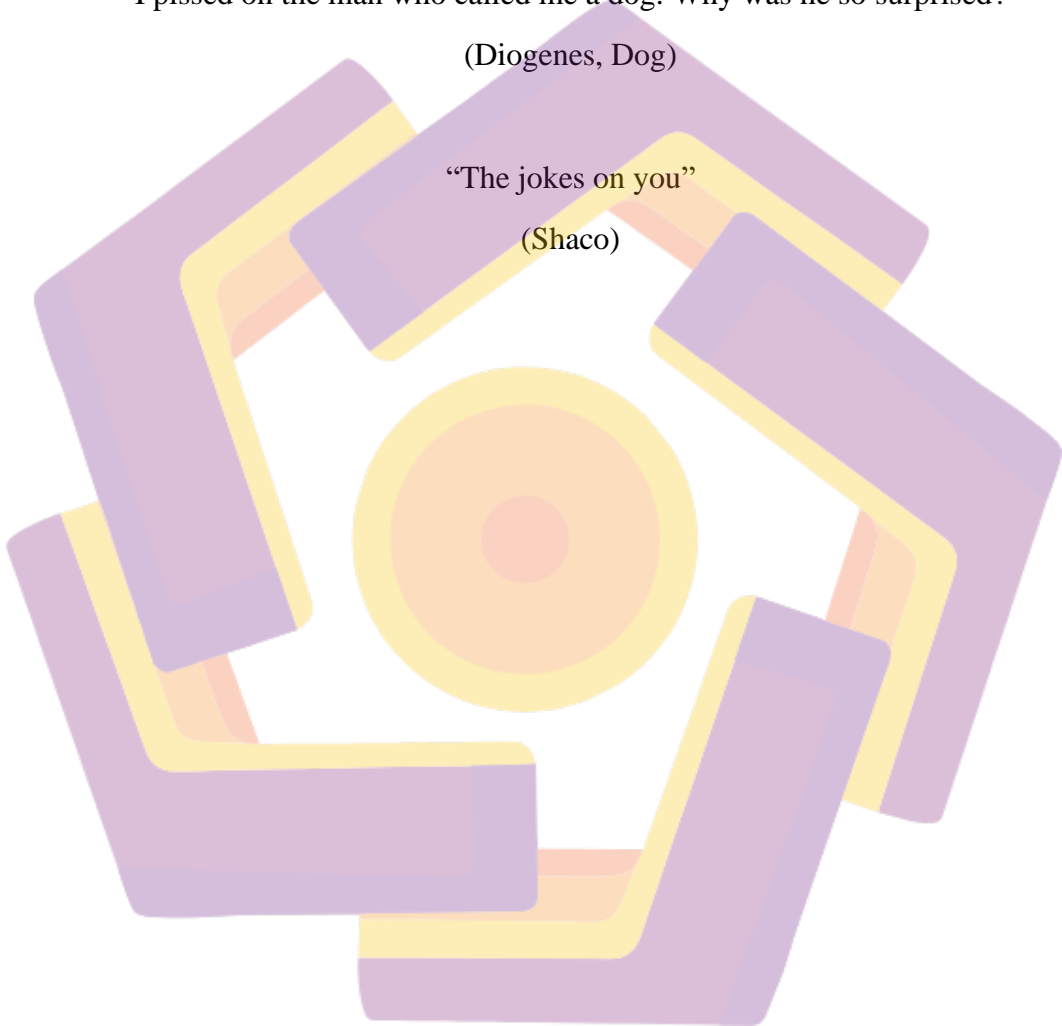
(JFK)

“I pissed on the man who called me a dog. Why was he so surprised?”

(Diogenes, Dog)

“The jokes on you”

(Shaco)



HALAMAN PERSEMBAHAN

Dengan segala puji dan syukur kepada Tuhan yang Maha Esa dan atas dukungan dan doa dari orang-orang tercinta, akhirnya skripsi ini dapat diselesaikan dengan baik. Oleh karena itu, dengan rasa bangga dan bahagia saya haturkan rasa syukur dan terimakasih saya kepada:

1. Tuhan Yang Maha Esa karena hanya atas izin dan karunia-Nyalah, maka skripsi ini dapat dibuat dan selesai pada waktunya.
2. Orang tua saya, yang tidak pernah lelah memberikan saya dukungan dan doa. Untuk orang tua yang tidak pernah lelah dalam memberikan semangat supaya saya bisa menyelesaikan skripsi ini yang telah banyak memberikan begitu banyak pengorbanan yang tidak bisa saya balas. Terimakasih banyak saya ucapkan untuk orang tua saya
3. Dosen Pembimbing skripsi bapak Kusnawi, S.Kom, M.Eng selaku dosen pembimbing saya, saya sangat berterimakasih atas bimbingannya selama ini yang telah memberikan masukan, kritik dan saran yang membangun agar menjadi lebih baik lagi untuk kedepannya. serta seluruh jajaran dosen Universitas Amikom Yogyakarta yang sudah membagikan ilmunya saya mengucapkan terimakasih, semoga ilmu dari bapak dan ibu dosen bisa saya amalkan ke yang lain juga.
4. Rekan – rekan kelas 18 Informatika 3, yang telah memberikan saya dukungan, semangat serta menemani selama 3 tahun dalam satu kelas yang penuh dengan segala kondisi dalam hidup. Terimakasih atas kenangan-kenangan yang telah kita ukir bersama-sama. Semoga kita menjadi orang-orang yang bermanfaat dan dikenang menjadi pribadi yang baik.

KATA PENGANTAR

Puji dan syukur penulis panjatkan kepada Tuhan Yang Maha Esa atas rahmat dan karunia-Nya yang sudah diberikan sehingga dapat menyelesaikan Karya Ilmiah Skripsi ini yang berjudul “Analisis Sentimen Animo Masyarakat Kabupaten Karanganyar Terhadap Vaksinisasi COVID-19 MENGGUNAKAN Algoritma SVM” Skripsi ini dapat terwujud berkat bimbingan, semangat, dan motivasi dari berbagai pihak, yaitu dari Dosen Pembimbing, Keluarga, Sahabat, dan Guru, oleh karena itu penulis ingin menyampaikan ucapan terimakasih yang sebanyakbanyaknya. Semoga bimbingan, semangat, dan motivasi yang telah diberikan menjadi suatu amal kebaikan dan semoga Tuhan memberikan balasan yang lebih baik dari apa yang telah diterima oleh penulis. Mengingat keterbatasan ilmu dan pengetahuan yang dimiliki, sehingga skripsi masih memiliki banyak kekurangan dan kelemahan. Besar harapan penulis untuk mendapatkan saran dan kritik dari pembaca. Semoga skripsi ini dapat bermanfaat dan berkontribusi bagi penulis dan masyarakat umum sebagai pembaca.

Yogyakarta, 15 Agustus 2023

Alfian Ady Chandra

DAFTAR ISI

HALAMAN JUDUL	ii
HALAMAN PERSETUJUAN	iii
HALAMAN PENGESAHAN.....	iv
HALAMAN PERNYATAAN KEASLIAN SKRIPSI.....	v
Halaman Pernyataan Keaslian Skripsi.....	v
MOTTO	vi
HALAMAN PERSEMBAHAN	vii
KATA PENGANTAR.....	viii
DAFTAR ISI.....	ix
DAFTAR TABEL	xii
DAFTAR GAMBAR.....	xiv
DAFTAR LAMBANG DAN SINGKATAN	xv
DAFTAR ISTILAH	xvi
INTISARI	xvii
Abstract.....	xviii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Perumusan masalah.....	2
1.3 Tujuan Penelitian.....	2
1.3 Batasan Masalah	3
1.4 Manfaat Penelitian	3
1.6.1 Metode Penelitian	3
1.6.1.1 Metode Pengumpulan Data.....	3
1.6.1.2 Studi Kepustakaan.....	4
1.6.1.3 Observasi.....	4
1.6.2 Metode Pengolahan Data.....	4
1.6.2.1 Cleaning.....	4
1.6.2.2 Case Folding.....	4
1.6.2.3 Tokenizing.....	4

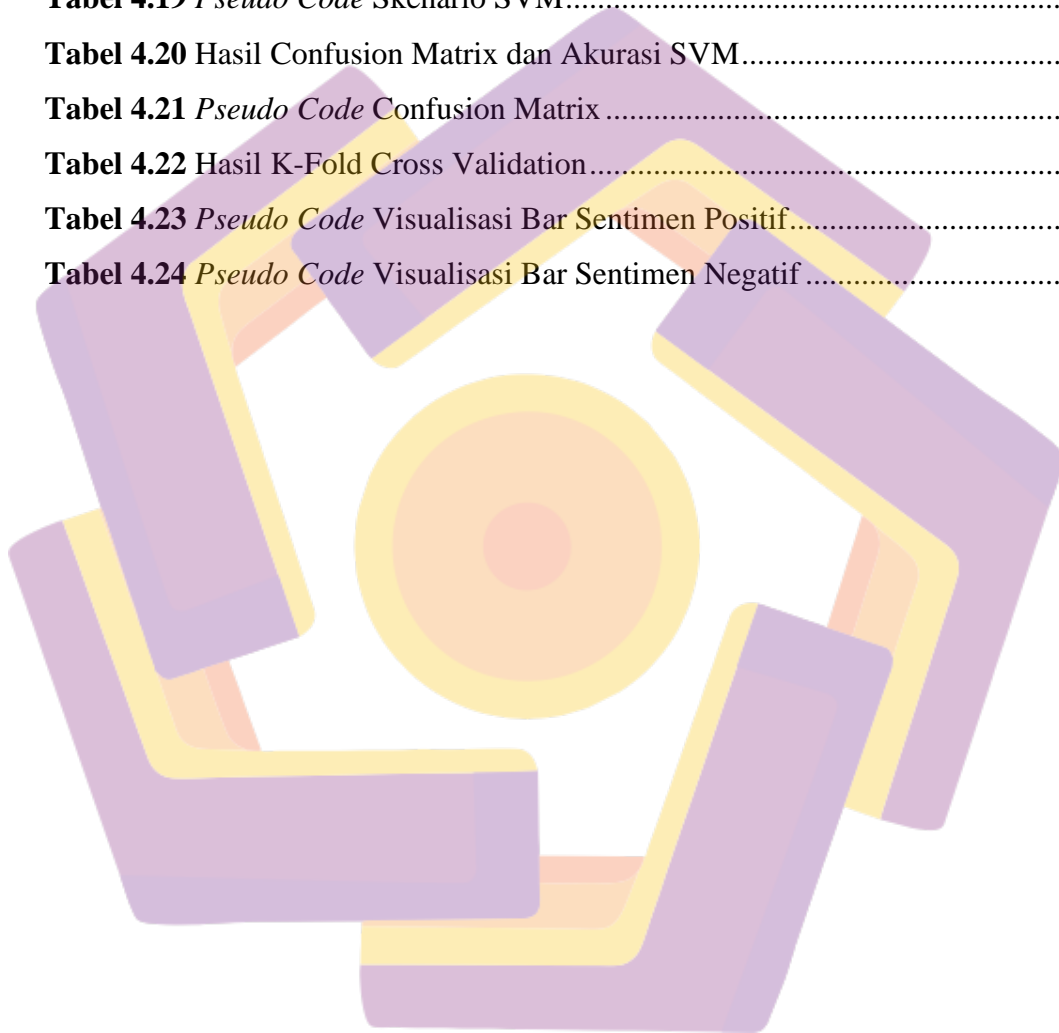
1.6.2.4 Filtering.....	4
1.6.2.5 Stopword.....	4
1.6.2.6 Stemming.....	5
1.6.3 Metode Analisis	5
1.6.4 Metode Pengujian.....	5
BAB II TINJAUAN PUSTAKA.....	6
2.1 Literature Review.....	6
2.2 Sentimen Analisis	12
2.3 Support Vector Machine.....	12
2.4 Twitter.....	15
2.5 Vaksinisasi	15
2.6 Covid-19	16
2.7 Data Mining	16
2.8 Web Scrapping.....	18
2.9 Preprocessing	19
2.9.1 Penentuan kelas atribut	19
2.9.2 Cleaning.....	19
2.9.3 Normalisasi Kalimat.....	20
2.9.4 Tokenisasi Kalimat.....	20
2.9.5 Stemming.....	20
2.10 Term Frequency-Inverse Document Frequency (TF-IDF).....	20
2.11 K-Fold Cross Validation.....	21
2.12 Confusion Matrix	22
BAB III METODOLOGI PENELITIAN	24
3.1 Alat dan Bahan.....	24
3.1.1 Alat.....	24
3.1.2 Bahan.....	24
3.2 Langkah Penelitian	24
3.3 Metode Pengumpulan Data	26
3.3.1 Studi Pustaka.....	26
3.3.2 Observasi.....	26
3.4 Metode Pengolahan dan Preprocessing.....	26

3.4.1 Cleaning.....	26
3.4.2 Case Folding.....	27
3.4.3 Stopword Removal.....	27
3.4.5 Tokenizing.....	28
3.4.6 Stemming.....	28
3.5 Metode Analisis Data	28
BAB IV HASIL DAN PEMBAHASAN	36
4.1 Crawling Data.....	36
4.2 Pengolahan Data.....	38
4.2.1 Preprocessing Data.....	39
4.2.2 Pelabelan.....	42
4.3 Klasifikasi dengan metode SVM	46
4.3.1 Data latih & data uji.....	46
4.3.2 TF-IDF.....	47
4.4 Hasil dan Evaluasi.....	49
4.5 Visualisasi.....	51
BAB V KESIMPULAN DAN SARAN	58
5.1 Kesimpulan.....	58
5.2 Saran.....	59
DAFTAR PUSTAKA	60

DAFTAR TABEL

Tabel 2.1 Penelitian yang pernah dilakukan	7
Tabel 2.2 Confusion Matrix	22
Tabel 3.1 Contoh Data Cleaning	27
Tabel 3.2 Contoh Case Folding	27
Tabel 3.3 Contoh Stopword Removal	27
Tabel 3.4 Contoh Tokenisasi	28
Tabel 3.5 Contoh Proses Stemming	28
Tabel 3.6 Contoh Data	29
Tabel 3.7 TF-IDF	30
Tabel 3.8 Matriks $K (x_i \cdot x_j^T)$	31
Tabel 3.9 bobot D1 sampai D4.....	32
Tabel 3.10 kemunculan <i>term</i> /kata pada kalimat baru.....	33
Tabel 3.11 Bobot <i>term</i> /kata baru pada data.....	34
Tabel 3.12 Matriks K hasil bobot yang diperoleh dari data baru	34
Tabel 3.13 Matriks Y kelas data yang dipakai menggunakan matriks Y	34
Tabel 3.14 Matriks x, x_i	35
Tabel 4.1 Contoh hasil <i>crawling</i> data.....	38
Tabel 4.2 Data Cleansing	39
Tabel 4.3 <i>Pseudo Code Data Cleaning</i>	39
Tabel 4.4 Contoh Data Removal Duplicate & na.....	40
Tabel 4.5 <i>Pseudo Code Data Cleaning</i>	40
Tabel 4.6 Contoh Stopword Removal	41
Tabel 4.7 <i>Pseudo Code Data Cleaning</i>	41
Tabel 4.8 Contoh Tokenizing	41
Tabel 4.9 <i>Pseudo Code Tokenizing</i>	42
Tabel 4.10 Hasil Pelabelan Kelas Sentimen.....	42
Tabel 4.11 <i>Pseudo Code Labelling</i>	43
Tabel 4.12 Hasil Labelling	44
Tabel 4.13 <i>Pseudo Code Pie Chart</i>	45

Tabel 4.14 Skenario Pembagian Data Uji	46
Tabel 4.15 <i>Pseudo Code</i> Pembagian data	47
Tabel 4.16 Kata Teratas Hasil pembobotan TF-IDF.....	47
Tabel 4.17 <i>Pseudo Code TF-IDF</i>	48
Tabel 4.18 Hasil Skenario SVM	49
Tabel 4.19 <i>Pseudo Code</i> Skenario SVM.....	49
Tabel 4.20 Hasil Confusion Matrix dan Akurasi SVM.....	50
Tabel 4.21 <i>Pseudo Code</i> Confusion Matrix	50
Tabel 4.22 Hasil K-Fold Cross Validation.....	51
Tabel 4.23 <i>Pseudo Code</i> Visualisasi Bar Sentimen Positif.....	52
Tabel 4.24 <i>Pseudo Code</i> Visualisasi Bar Sentimen Negatif	53



DAFTAR GAMBAR

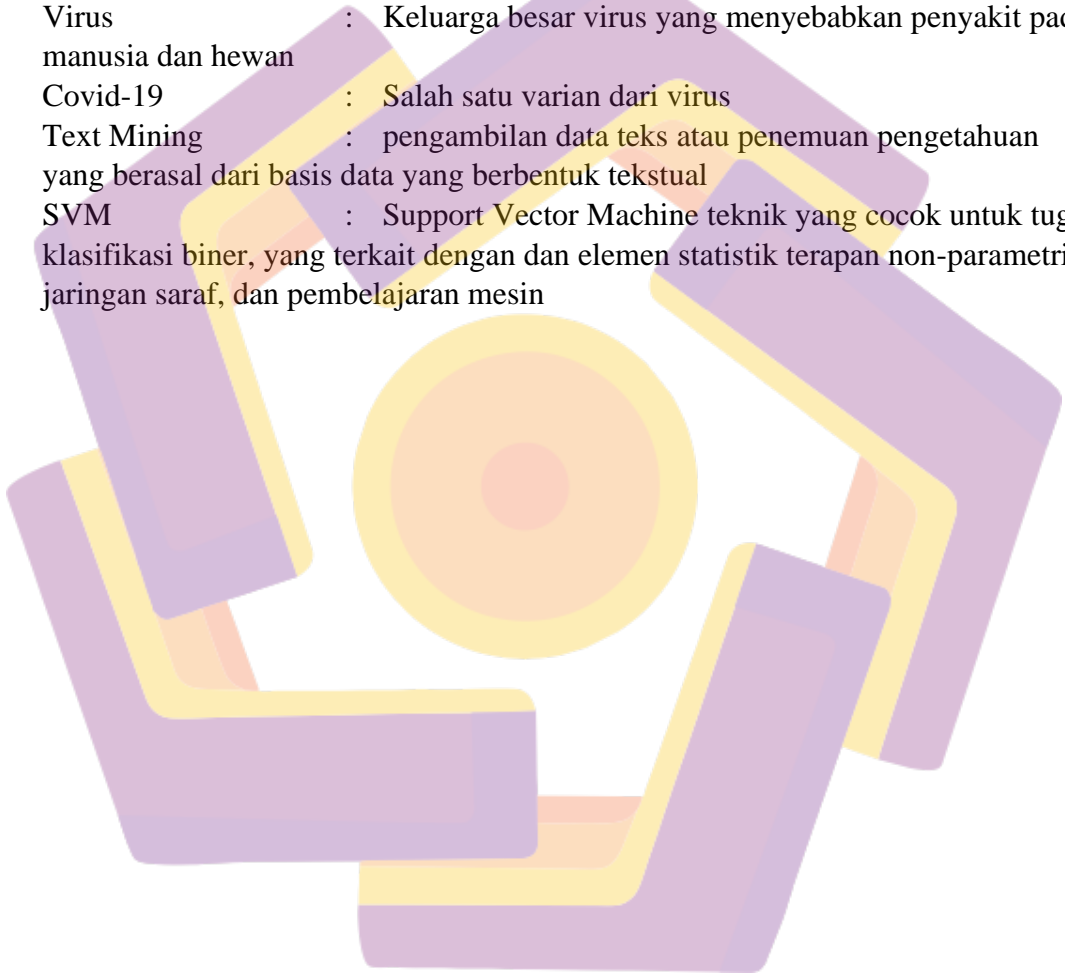
Gambar 2 1 Contoh SVM Hyperplane (Sumber Azizah, Syuradi,2019).....	13
Gambar 2 2 Contoh <i>One Against One</i>	14
Gambar 2 3 Contoh <i>One Against All</i>	15
Gambar 2 4 Contoh tahapan data mining.....	17
Gambar 3.1 Alur Penelitian	25
Gambar 4.1 Anaconda Navigator	36
Gambar 4. 2 Jupyter Lab.....	36
Gambar 4.3 Crawling Data	37
Gambar 4.4 Pie Chart hasil labelling	45
Gambar 4.5 Kata yang sering muncul (positif).....	54
Gambar 4.6 Wordcloud (positif).....	55
Gambar 4. 7 Kata yang sering muncul (negatif).....	56
Gambar 4. 8 Worcloud (negatif).....	56

DAFTAR LAMBANG DAN SINGKATAN

W_{ij}	: Menunjukkan bobot kata/term t_j terhadap dokumen d_i
tf_{ij}	: Jumlah kemunculan kata/term t_j dalam dokumen d_i
Idf	: <i>Inverse Document Frequency</i>
D	: Jumlah dokumen
df_i	: Jumlah kemunculan term terhadap D
$P(T_n C)$: Kondisi probabilitas kata ke- n dengan diketahui kelas c
W_{ct}	: Nilai pembobotan TF-IDF dari term- t kelas c
$\sum W'_{ct}$: Jumlah total TF-IDF seluruh term pada kelas c .
B	: Jumlah kata unik pada seluruh dokumen
$\sum_{i=1}^e$: Cara untuk menuliskan penjumlahan beruntun secara singkat dengan batas nilai bawah i adalah 1 dan batas nilai atas adalah e
α_i	: Alfa adalah nilai yang dijadikan sebagai tolak ukur untuk menentukan taraf kepercayaan atau generalisasi dari objek yang diteliti setelah dilakukan analisa dan interpretasi data.
x_i^T	: Matriks persegi dengan hasil transposnya berupa dirinya sendiri

DAFTAR ISTILAH

- Analisis Sentimen : Sebuah kecerdasan buatan yang berdasarkan tentang sebuah opini maupun emosi
- NLP : Natural Language Processing yang berarti analisis pada teks, komputasi linguistik, dan pembelajaran mesin sebagai kecerdasan buatan untuk menganalisis bahasa tertulis orang secara sistematis
- Virus : Keluarga besar virus yang menyebabkan penyakit pada manusia dan hewan
- Covid-19 : Salah satu varian dari virus
- Text Mining : pengambilan data teks atau penemuan pengetahuan yang berasal dari basis data yang berbentuk tekstual
- SVM : Support Vector Machine teknik yang cocok untuk tugas klasifikasi biner, yang terkait dengan dan elemen statistik terapan non-parametrik, jaringan saraf, dan pembelajaran mesin



INTISARI

Virus Covid-19 merupakan keluarga besar virus yang menyebabkan penyakit pada manusia dan hewan. Penyakit ini termasuk penyakit menular dan menyerang pernafasan. Virus ini berasal dari Negara China, lebih tepatnya di Kota Wuhan, Provinsi Hubei. Hingga pada awal bulan Maret tahun 2020 virus ini masuk ke Indonesia. Pemerintah berupaya menangani masalah ini dengan melakukan vaksinasi Covid-19 secara menyeluruh secara bertahap.

Dengan adanya vaksinasi Covid-19 secara menyeluruh di Indonesia, tentu saja tidak luput dengan adanya berbagai polemik yang terjadi di tengah masyarakat kita,. Penyebaran vaksinasi memang sudah menyeluruh, tetapi tentu saja masih ada warga masyarakat khususnya Kabupaten Karanganyar sendiri yang masih ragu untuk melakukan vaksinasi Covid-19 dengan berbagai pendapat mereka sendiri. Maka dari itu, akan dilakukan sebuah analisis sentiment terhadap masyarakat Kabupaten Karanganyar menggunakan metode SVM (Support Vector Machine) untuk menemukan sebuah data yang valid tentang bagaimana animo yang sedang terjadi di masyarakat Kabupaten Karanganyar tentang dilakukannya vaksinasi Covid-19.

Berdasarkan latar belakang tersebut, maka penelitian ini dapat dirumuskan bahwa bagaimana respon dan opini masyarakat Kabupaten Karanganyar terhadap vaksinasi COVID-19 menggunakan data yang diambil dari media sosial twitter. Algoritma yang digunakan pada penelitian ini adalah SVM. Algoritma SVM dipilih karena mempunyai kelebihan sederhana dan akurasi yang tinggi. Tujuan dari penelitian ini adalah mengklasifikasikan sentimen positif dan negatif dari twitter terhadap kata kunci yaitu “Vaksin” dan “Covid”. Hasil penelitian menggunakan algoritma SVM mendapatkan nilai akurasi dengan metode pengujian confusion matrix sebesar 64% dan k-fold cross validation sebesar 60,8%..

Kata Kunci: Analisis sentimen, vaksin, covid, svm

Abstract

The Covid-19 virus is a large family of viruses that cause disease in humans and animals. This disease includes infectious and respiratory diseases. This virus originated in China, more precisely in Wuhan City, Hubei Province. Until the beginning of March 2020 this virus entered Indonesia. The government is trying to deal with this problem by vaccinating the Covid-19 as a whole in stages.

With the comprehensive Covid-19 vaccination in Indonesia, of course, there are various polemics that occur in our society. The spread of vaccination has indeed been comprehensive, but of course there are still residents, especially Karanganyar Regency itself, who are still hesitant to vaccinate Covid-19 with their own opinions. Therefore, a sentiment analysis will be carried out on the people of Karanganyar Regency using the SVM (Support Vector Machine) method to find valid data about how the enthusiasm that is going on in the Karanganyar Regency community regarding the Covid-19 vaccination will be carried out.

Based on this background, this research can be formulated that how is the response and opinion of the people of Karanganyar Regency to the COVID-19 vaccination using data taken from Twitter social media. The algorithm used in this study is SVM. The SVM algorithm was chosen because it has the advantages of simplicity and high accuracy. The purpose of this study is to classify positive and negative sentiments from Twitter on the keywords "Vaccine" and "Covid". The results of the research using the SVM algorithm get an accuracy value with the confusion matrix testing method of 64% and k-fold cross validation of 60.8%.

Keyword: *Sentiment analysis, vaccine, covid, svm*