



**SENTIMEN ANALISIS MENGENAI POLUSI UDARA  
MENGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE DAN  
RANDOM FOREST**

**LAPORAN SKRIPSI**

**MUHAMMAD VARIANSJAH DALIMUNTHE**

**41519010191**

**UNIVERSITAS**

**MERCU BUANA**

**Diajukan sebagai salah satu syarat untuk memperoleh gelar sarjana**

**PROGRAM STUDI TEKNIK INFORMATIKA**

**FAKULTAS ILMU KOMPUTER**

**UNIVERSITAS MERCU BUANA JAKARTA**

**2023**

## HALAMAN PERNYATAAN KARYA SENDIRI

Saya yang bertanda tangan di bawah ini:

Nama : Muhammad Variansjah  
NIM : 41519010191  
Program Studi : Teknik Informatika  
Judul Laporan Skripsi : Sentimen Analisis Mengenai Polusi Udara  
Menggunakan Algoritma *Support Vector Machine* dan *Random Forest*

Menyatakan bahwa Laporan Skripsi ini adalah hasil karya saya sendiri dan bukan plagiat, serta semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar. Apabila ternyata ditemukan di dalam Laporan Skripsi saya terdapat unsur plagiat, maka saya siap mendapatkan sanksi akademis yang berlaku di Universitas Mercu Buana.

UNIVERSITAS  
MERCU BUANA

Jakarta, 22 Desember 2023



Muhammad Variansjah Dalimunthe

## HALAMAN PENGESAHAN

Laporan Skripsi ini diajukan oleh:

Nama : Muhammad Variansjah Dalimunthe  
NIM : 41519010191  
Program Studi : Teknik Informatika  
Judul Laporan Skripsi : Sentimen Analisis Mengenai Polusi Udara Menggunakan  
Algoritma *Support Vector Machine* dan *Random Forest*

Telah berhasil dipertahankan pada sidang di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Strata 1 pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer Universitas Mercu Buana.

Disahkan oleh:

Pembimbing : Lukman Hakim, ST, M.Kom  
NIDN : 0327107701



Ketua Penguji : Dr. Hadi Santoso, S.Kom., M.Kom  
NIDN : 0225067701



Penguji 1 : Harni Kusniyati, ST., M.Kom  
NIDN : 0324068101



Penguji 2 : Dwi Anindyani Rochmah, ST,MTI  
NIDN : 0011057801



UNIVERSITAS  
MERCU BUANA

Jakarta, 22 Desember 2023

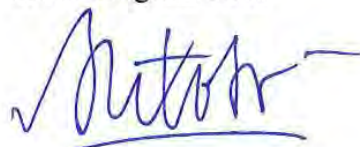
Mengetahui,

Dekan

Ketua Program Studi



Dr. Bambang Jekonowo, S.Si., MTI



Dr. Hadi Santoso, S.Kom., M.Kom

## KATA PENGANTAR

Puji syukur kehadirat Tuhan yang Maha Esa, atas segala rahmat dan ridha-Nya sehingga penulis dapat menyelesaikan tugas akhir yang merupakan salah satu persyaratan kelulusan Program Studi Strata Satu (S1) pada jurusan Teknik Informatika, Universitas Mercu Buana.

Penulis menyadari bahwa tugas akhir ini masih jauh dari sempurna, karena kesempurnaan sejatinya hanya milik Tuhan yang Maha Esa. Oleh karena itu, saran dan masukan yang membangun senantiasa penulis terima dengan senang hati. Serta berkat dukungan, motivasi, bantuan, bimbingan, dan doa dari banyak pihak, penulis mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. Andi Adriansyah, M.Eng. selaku Rektor Universitas Mercu Buana.
2. Bapak Yaya Sudarya Triana, M.Kom., Ph.D. selaku Dekan Fakultas Ilmu Komputer.
3. Bapak Ir. Emil R. Kaburuan, Ph.D., IPM. selaku Ketua Program Studi Teknik Informatika Universitas Mercubuana.
4. Ibu Dosen Pembimbing, S.Kom., M.Kom. selaku dosen pembimbing tugas akhir yang telah memberikan pengarahan, motivasi, menyediakan waktu, tenaga, dan pikiran sehingga selama pembuatan tugas akhir ini terjadwal dengan baik.
5. Kedua Orang Tua saya yang selalu mensupport dan mendukung saya selama menjalani masa studi sebagai mahasiswa Universitas Mercubuana..
6. Semua teman kuliah yang selalu berbagi informasi dan memberikan dukungan dalam bentuk yang berbeda-beda.

Akhir kata, penulis berharap semoga Tuhan yang Maha Esa membalas kebaikan dan selalu mencurahkan rahmat, hidayah, serta panjang umur kepada kita semua, aamiin. Terima Kasih.

Jakarta, 22 Desember 2023

Muhammad Variansjah Dalimunthe

## HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademik Universitas Mercu Buana, saya yang bertanda tangan di bawah ini:

Nama : Muhammad Variansjah Dalimunthe  
NIM : 41519010191  
Program Studi : Teknik Informatika  
Judul Laporan Skripsi : Sentimen Analisis Mengenai Polusi Udara Menggunakan  
Algoritma *Support Vector Machine* dan *Random Forest*

Demi pengembangan ilmu pengetahuan, dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Non-Eksklusif** (*Non-exclusive Royalty-Free Right*) atas karya ilmiah saya yang berjudul di atas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti Non-Eksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan Laporan Magang/Skripsi/Tesis/Disertasi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 22 Desember 2023

Yang menyatakan,



Muhammad Variansjah Dalimunthe

UNIVERSITAS  
MERCU BUANA

BERBAHASA INDONESIA

## ABSTRAK

Nama : Muhammad Variansjah  
NIM : 41519010191  
Program Studi : Teknik Informatika  
Judul Laporan Skripsi : Sentimen Analisis Mengenai Polusi Udara Menggunakan Algoritma *Support Vector Machine* dan *Random Forest*  
Pembimbing : Lukman Hakim, ST, M.Kom

Polusi udara adalah kontaminasi area dalam dan luar ruangan oleh zat kimia, fisik, atau biologis yang mengubah sifat alami atmosfer. Insinerator domestik, mobil, pabrik, dan kebakaran hutan merupakan sumber polusi udara yang umum. Di Indonesia, tidak diragukan lagi kalau polusi udara terjadi karena banyaknya kebakaran hutan di Indonesia. Akibat kasus tersebut, banyak opini masyarakat yang berbeda-beda. Berbagai sentiment terjadi di dunia maya, salah satunya *Twitter*. *Twitter* adalah *social media* yang paling banyak menampung berbagai macam opini positif, negatif maupun netral. Oleh karena itu, peneliti ingin memecahkan masalah dengan implementasi algoritma SVM dan *Random Forest*. *Dataset* didapatkan dari hasil scrapping menggunakan *tweet harvest*. Data yang diperoleh didapatkan sebanyak 8555 *tweet*. Dengan membagi model *dataset* 80% dan 20%, hasil didapat bahwa akurasi algoritma SVM lebih baik dari algoritma *Random Forest*. Akurasi dari algoritma SVM sebesar 83% sedangkan algoritma *Random Forest* sebesar 81%.

**Katakunci : *sentiment analisis, SVM, Random Forest, Twitter, Polusi***

BERBAHASA INGGRIS

## ABSTRACT

Nama : Muhammad Variansjah  
NIM : 41519010191  
Program Studi : Teknik Informatika  
Judul Laporan Skripsi : Sentimen Analisis Mengenai Polusi Udara Menggunakan Algoritma *Support Vector Machine* dan *Random Forest*  
Pembimbing : Lukman Hakim, ST, M.Kom

*The term "air pollution" refers to the contaminating effects of chemical, physical, or biological contaminants on the atmosphere that alter its inherent characteristics in both indoor and outdoor spaces. Common sources of air pollution include automobiles, factories, homes with incinerators, and forest fires. There is no denying that Indonesia's numerous forest fires contribute to air pollution there. This situation has caused divisions in the opinions of a lot of people. Twitter is only one among the numerous diverse emotions that can be felt on the internet. Twitter is the social media site that allows the most variety of unbiased, affirmative, and negative opinions. Therefore, the research's objective is to solve the problem by using the Random Forest and SVM methods. Tweet Harvest was used to scrape results and compile the information. The data collection contained 8555 tweets. by dividing.*

**Keywords:** *sentiment analisis, SVM, Random Forest, Twitter, Polusi*

## DAFTAR ISI

<b>HALAMAN JUDUL .....</b>	<b>i</b>
<b>HALAMAN PERNYATAAN KARYA SENDIRI .....</b>	<b>ii</b>
<b>HALAMAN PENGESAHAN.....</b>	<b>iii</b>
<b>KATA PENGANTAR.....</b>	<b>iv</b>
<b>HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS.....</b>	<b>v</b>
<b>ABSTRAK .....</b>	<b>vi</b>
<b>ABSTRACT .....</b>	<b>vii</b>
<b>DAFTAR ISI.....</b>	<b>viii</b>
<b>DAFTAR TABEL .....</b>	<b>x</b>
<b>DAFTAR GAMBAR.....</b>	<b>xi</b>
<b>DAFTAR LAMPIRAN .....</b>	<b>xii</b>
<b>BAB I PENDAHULUAN.....</b>	<b>1</b>
1.1. Latar Belakang .....	1
1.2. Perumusan Masalah.....	2
1.3. Tujuan Penelitian.....	2
1.4. Manfaat Penelitian.....	2
1.5. Batasan Penelitian .....	2
<b>BAB II TINJAUAN PUSTAKA .....</b>	<b>4</b>
2.1. Penelitian Terdahulu .....	4
Analisis Sentimen Algoritma SVM pada Data <i>Twitter</i> Komisi Pemberantasan Korupsi Republik Indonesia .....	7
2.2. Teori Pendukung .....	9
<b>BAB III METODE PENELITIAN .....</b>	<b>12</b>
3.1. Jenis Penelitian.....	12
3.2. Metode Pengumpulan Data .....	12
3.3. Tahapan Penelitian .....	12
<b>BAB IV HASIL DAN PEMBAHASAN .....</b>	<b>16</b>
4.1. Dataset.....	16
4.2. <i>Pre-Processing</i> .....	17
4.2.1. Proses Cleaning.....	17
4.2.2. Proses Case Folding .....	18



4.2.3.	Proses Normalisasi.....	18
4.2.4.	Proses TF-IDF.....	31
4.3.	Pembuatan Model.....	31
4.4.	Visualisasi Data.....	32
4.5.	Implementasi Algoritma.....	33
4.6.	Evaluasi dan Perbandingan Algoritma.....	35
<b>BAB V</b>	<b>KESIMPULAN DAN SARAN.....</b>	<b>37</b>
5.1	Kesimpulan.....	37
5.2	Saran.....	37
<b>Daftar Pustaka</b>	<b>.....</b>	<b>38</b>
<b>LAMPIRAN</b>	<b>.....</b>	<b>40</b>



## DAFTAR TABEL

Tabel 1. Penelitian Terdahulu.....	9
Tabel 2. Proses Cleaning .....	18
Tabel 3. Proses Normalisasi.....	21
Tabel 4. Proses Stopwords Removal .....	24
Tabel 5. Proses Tokenisasi.....	27
Tabel 6. Proses Labeling.....	31



UNIVERSITAS  
MERCU BUANA

## DAFTAR GAMBAR

Gambar 1. Flowchart Penelitian .....	12
Gambar 2. Sample Data Mentah.....	16
Gambar 3. Pembagian Model Uji dan Model Training .....	31
Gambar 4. Visualisasi Pychart Dataset.....	32
Gambar 5. Visualisasi Wordcloud Dataset Keseluruhan.....	33
Gambar 6. Visualisasi Wordcloud Dataset Positif dan Negatif.....	33
Gambar 7. Hasil Performa Algoritma Random Forest .....	34
Gambar 8. Hasil Performa Algoritma SVM .....	34
Gambar 9. Confusion Matrix SVM .....	34
Gambar 10. Confusion Matrix Random Forest .....	34
Gambar 11. Visualisasi Bar Plot.....	35
Gambar 12. Hasil Perbandingan Algoritma SVM dan Random Forest.....	35



## DAFTAR LAMPIRAN

Lampiran 1. Kartu Asistensi Tugas Akhir .....	40
Lampiran 2. Bukti Submit Jurnal.....	41
Lampiran 3. Naskah Jurnal Yang Disubmit.....	48
Lampiran 4. Surat Persetujuan.....	49
Lampiran 5. Curriculum Vitae.....	51
Lampiran 6. Plagiarisme Check.....	52
Lampiran 7. Lampiran Surat HKI.....	54
Lampiran 8. Bukti Sertifikat BNSP .....	55

