

**PENERAPAN ALGORITMA *SUPPORT VECTOR MACHINE*
UNTUK MENDETEKSI EMOSI DARI TEKS BAHASA
INDONESIA**

SKRIPSI



BHAYU AJI SENO

5235153266

Skripsi ini Ditulis untuk Memenuhi Sebagian Persyaratan dalam Memperoleh
Gelar Sarjana

Mencerdaskan dan
PENDIDIKAN TEKNIK INFORMATIKA DAN KOMPUTER

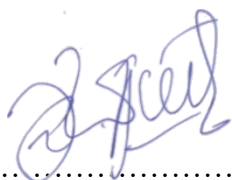

FAKULTAS TEKNIK
UNIVERSITAS NEGERI JAKARTA

2023

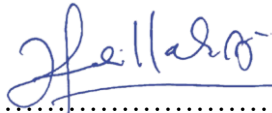

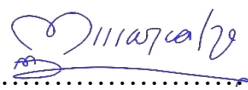
**PENERAPAN ALGORITMA *SUPPORT VECTOR MACHINE* UNTUK
MENDETEKSI EMOSI DARI TEKS BAHASA INDONESIA**

Nama: Bhayu Aji Seno, NIM: 5235153266

HALAMAN PENGESAHAN

NAMA DOSEN	TANDA TANGAN	TANGGAL
Dr. Widodo, S.Kom, M.Kom. (Dosen Pembimbing 1)		13-02-2023
Bambang Prasetya Adhi, M.Kom. (Dosen Pembimbing 2)		13-02-2023

PENGESAHAN PANITIA UJIAN SKRIPSI

NAMA DOSEN	TANDA TANGAN	TANGGAL
Hamidillah Ajie, S.Si, M.T. (Ketua Penguji)		10-02-2023
Diat Nurhidayat, M.T.I. (Dosen Penguji 1)		10-02-2023
Murien Nugraheni, S.T., M.Cs. (Dosen Penguji 2)		13-02-2023

Tanggal Lulus :

HALAMAN PERNYATAAN

Dengan ini peneliti menyatakan bahwa :

1. Karya tulis skripsi yang berjudul Penerapan Algoritma *Support Vector Machine* untuk Mendeteksi Emosi dari Teks Bahasa Indonesia adalah asli dan belum pernah diajukan untuk mendapatkan gelar akademik sarjana, baik di Universitas Negeri Jakarta maupun di perguruan tinggi lainnya;
2. Karya tulis skripsi yang berjudul Penerapan Algoritma *Support Vector Machine* untuk Mendeteksi Emosi dari Teks Bahasa Indonesia adalah murni gagasan, rumusan, dan penelitian peneliti sendiri dengan arahan dari dosen pembimbing;
3. Dalam karya tulis, tidak terdapat karya atau pendapat yang telah ditulis atau dipublikasi orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan dicantumkan dalam daftar pustaka;
4. Pernyataan peneliti dibuat dengan sesungguhnya dan apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka peneliti bersedia menerima sanksi akademik berupa pencabutan gelar yang telah diperoleh karena karya tulis ini, serta sanksi lainnya sesuai norma yang berlaku di Universitas Negeri Jakarta.

Jakarta, 27 Januari 2023

Yang Membuat Pernyataan



Bhayu Aji Seno

5235153266



KEMENTERIAN PENDIDIKAN DAN KEBUDAYAAN
UNIVERSITAS NEGERI JAKARTA
UPT PERPUSTAKAAN

Jalan Rawamangun Muka Jakarta 13220
Telepon/Faksimili: 021-4894221
Laman: lib.unj.ac.id

**LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS**

Sebagai sivitas akademika Universitas Negeri Jakarta, yang bertanda tangan di bawah ini, saya:

Nama : BHAYU AJI SENO
NIM : 5235153266
Fakultas/Prodi : Teknik / Pendidikan Teknik Informatika dan Komputer
Alamat email : bhayuajiseno@gmail.com

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada UPT Perpustakaan Universitas Negeri Jakarta, Hak Bebas Royalti Non-Eksklusif atas karya ilmiah:

Skripsi Tesis Disertasi Lain-lain (.....)

yang berjudul :

Penerapan Algoritma *Support Vector Machine* untuk Mendeteksi Emosi dari Teks

Bahasa Indonesia

Dengan Hak Bebas Royalti Non-Eksklusif ini UPT Perpustakaan Universitas Negeri Jakarta berhak menyimpan, mengalihmediakan, mengelolanya dalam bentuk pangkalan data (*database*), mendistribusikannya, dan menampilkan/mempublikasikannya di internet atau media lain secara *fulltext* untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan atau penerbit yang bersangkutan.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Perpustakaan Universitas Negeri Jakarta, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 1 Maret 2023

Penulis

(BHAYU AJI SENO)

nama dan tanda tangan

KATA PENGANTAR

Puji dan syukur penulis panjatkan ke hadirat Allah SWT yang telah memberikan rahmat serta nikmat-Nya dan tak lupa shalawat serta salam senantiasa dicurahkan kepada Nabi Muhammad SAW yang telah menunjukkan jalan ke zaman yang terang benderang. Berdasarkan hal tersebut penulis dapat menyelesaikan skripsi dengan judul “Penerapan Algoritma *Support Vector Machine* untuk Mendeteksi Emosi dari Teks Bahasa Indonesia” yang merupakan persyaratan untuk meraih gelar Sarjana Pendidikan.

Skripsi ini dapat terwujud dengan baik berkat bimbingan, dorongan, serta saran-saran, dan bantuan dari berbagai pihak. Pada kesempatan ini, penulis ingin menyampaikan ucapan terimakasih yang sebesar-besarnya kepada:

1. Bapak Dr. Widodo, S.Kom, M.Kom. selaku Ketua Program Studi Pendidikan Teknik Informatika dan Komputer Fakultas Teknik Universitas Negeri Jakarta;
2. Bapak Dr. Widodo, S.Kom, M.Kom. selaku dosen pembimbing I dan Bapak Bambang Prasetya Adhi, M.Kom. selaku dosen pembimbing II yang telah memberikan arahan kepada penulis dalam menyelesaikan skripsi ini;
3. Kedua orang tua, (Alm) Bapak Suwandi Amin dan (Alm) Ibu Siti Djariah yang menjadi sumber motivasi penulis dalam menyelesaikan skripsi ini;
4. Isnaeni Muthmainah yang membantu penulis dalam menyelesaikan skripsi ini;
5. Teman-teman PTIK 2015 UNJ yang senantiasa memberikan saran, pengalaman, semangat, dan do'a dalam penyusunan skripsi ini;
6. Seluruh pihak yang terlibat secara langsung maupun tidak langsung yang tidak bisa disebutkan satu per satu.

Penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan, maka dari itu penulis memohon maaf apabila terdapat kekurangan dan kesalahan, baik dari isi maupun tulisan. Akhir kata penulis berharap agar penulisan dan penyusunan skripsi ini dapat bermanfaat bagi para pembaca dan semua pihak terkait.

Jakarta, 27 Januari 2023

Bhayu Aji Seno

ABSTRAK

BHAYU AJI SENO. Penerapan Algoritma *Support Vector Machine* untuk Mendeteksi Emosi dari Teks Bahasa Indonesia. Skripsi, Program Studi Pendidikan Teknik Informatika dan Komputer, Fakultas Teknik, Universitas Negeri Jakarta. Januari 2023.

Pertumbuhan penggunaan internet khususnya sosial media yang sangat pesat telah menyebabkan perubahan interaksi antar manusia yang sebelumnya terjadi secara langsung di dunia nyata kemudian berubah menjadi tidak langsung melalui sosial media. Ketika manusia berinteraksi melalui sosial media, terkadang terdapat kesalahpahaman informasi yang didapatkan dibandingkan secara langsung di dunia nyata. Hal ini karena kurangnya faktor penting ketika manusia berinteraksi melalui sosial media, yaitu emosi. Penelitian ini bertujuan untuk mendeteksi emosi dari teks bahasa Indonesia dengan menerapkan algoritma *Support Vector Machine* (SVM). Penelitian dimulai dengan pembuatan 4 jenis data sampel yaitu, data sampel dengan jumlah 1000, 2500, dan 4000 tweet untuk proses *training*, dan data sampel dengan jumlah 120 tweet untuk proses *testing*. Kemudian seluruh data sampel tersebut melalui proses *Natural Leanguage Processing* (NLP) untuk diekstrak informasinya. Setelah itu dilakukan proses training dengan algoritma SVM menggunakan 3 tipe kernel yang berbeda, yaitu Polinomial, Gaussian RBF, dan Sigmoid. Kemudian dilakukan pengujian dengan hasil tingkat akurasi yang cukup variatif untuk setiap jenis data sampel serta tipe kernel yang digunakan. Hasil penelitian mendapatkan kesimpulan semakin banyak jumlah data yang digunakan, maka semakin tinggi tingkat akurasi yang dihasilkan, namun waktu proses *training* yang dibutuhkan lebih lama. Sementara untuk tipe kernel Sigmoid memiliki tingkat akurasi yang tinggi, serta waktu proses *training* yang lebih cepat dibandingkan tipe kernel lainnya.

Kata kunci: *Support Vector Machine*, *Natural Leanguage Processing*, Emosi

*Mencerdaskan dan
Memartabatkan Bangsa*

ABSTRACT

BHAYU AJI SENO. *Application of Support Vector Machine Algorithm to Detect Emotion From Indonesian Text. Mini Thesis, Informatics and Computer Engineering Education Study Program, Faculty of Engineering, Universitas Negeri Jakarta. January 2023.*

The rapid growth of internet use, especially social media, has led to changes in interactions between humans that previously occurred directly in the real world and then changed to become indirect through social media. When humans interact through social media, sometimes there is a misunderstanding of the information obtained compared directly to the real world. This is due to the lack of important factors when humans interact through social media, namely emotions. This study aims to detect emotions from Indonesian text by applying the Support Vector Machine (SVM) algorithm. The research began by making 4 types of sample data, namely sample data with a total of 1000, 2500 and 4000 tweets for the training process, and sample data with a total of 120 tweets for the testing process. Then all the sample data went through the Natural Language Processing (NLP) process to extract the information. After that, the training process was carried out with the SVM algorithm using 3 different kernel types, namely Polynomial, Gaussian RBF, and Sigmoid. Then testing was carried out with the results of a fairly varied level of accuracy for each type of sample data and the type of kernel used. The results of the study concluded that more amount of data used, the higher level of accuracy produced, but the training process time required is longer. Meanwhile, the Sigmoid kernel type has a high level of accuracy, as well as a faster training processing time than other kernel types.

Keywords: Support Vector Machine, Natural Language Processing, Emotions

*Mencerdaskan dan
Memartabatkan Bangsa*

DAFTAR ISI

KATA PENGANTAR	iv
DAFTAR ISI.....	vii
DAFTAR TABEL.....	ix
DAFTAR GAMBAR	x
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang Masalah.....	1
1.2 Identifikasi Masalah.....	3
1.3 Batasan Masalah.....	3
1.4 Rumusan Masalah.....	4
1.5 Tujuan Penelitian	4
1.6 Manfaat Penelitian	4
BAB II TINJAUAN PUSTAKA.....	5
2.1 Teori	5
2.1.1 Model Emosi	5
2.1.2 <i>Affective Computing</i>	6
2.1.3 Deteksi Emosi	6
2.1.4 Pendekatan <i>Learning Based</i>	7
2.1.5 <i>Natural Language Processing (NLP)</i>	8
2.1.5.1 <i>Text Preprocessing</i>	9
2.1.5.2 <i>Feature Extraction</i>	10
2.1.6 <i>Support Vector Machine (SVM)</i>	10
2.1.6.1 SVM pada <i>Linearly Separable Data</i>	12
2.1.6.2 <i>Hyperplane SVM</i>	14
2.1.6.3 SVM pada <i>Non-Linearly Separable Data</i>	17
2.2 Penelitian Relevan.....	20
2.3 Kerangka Berpikir.....	23
BAB III METODOLOGI PENELITIAN.....	26
3.1 Tempat dan Waktu Penelitian.....	26
3.2 Alat dan Bahan Penelitian.....	26



3.2.1	Alat Penelitian	26
3.2.2	Bahan Penelitian	26
3.3	Diagram Alir Penelitian	27
3.3.1	Membuat Jenis Data Sampel	28
3.3.2	Mengekstrak Informasi dengan NLP	28
3.3.3	Melatih Algoritma SVM	29
3.3.4	Menguji Algoritma SVM	29
3.3.5	Mengevaluasi Algoritma SVM	30
BAB IV HASIL PENELITIAN		31
4.1	Deskripsi Hasil Penelitian	31
4.2	Hasil Penelitian	31
4.2.1	Membuat Data Sampel	31
4.2.2	Mengekstrak Informasi dengan NLP	32
4.2.3	Melatih Algoritma SVM	32
4.2.4	Menguji Algoritma SVM	33
4.3	Pembahasan	33
BAB V PENUTUP		36
5.1	Kesimpulan	36
5.2	Saran	36
DAFTAR PUSTAKA		37
LAMPIRAN-LAMPIRAN		39
DAFTAR RIWAYAT HIDUP		46

*Mencerdaskan dan
Memartabatkan Bangsa*

DAFTAR TABEL

Tabel 2.1. Ilustrasi Empat Kategori <i>Affective Computing</i>	6
Tabel 2.2. Fungsi Kernel.....	19
Tabel 2.3. Rangkuman Penelitian Relevan	22
Tabel 4.1. Perbandingan Waktu Proses NLP per Jenis Data Sampel	32
Tabel 4.2. Perbandingan Waktu Proses <i>Training</i> per Jenis Data Sampel dan Tipe Kernel.....	32
Tabel 4.3. Perbandingan Tingkat Akurasi Algoritma SVM per Jenis Data Sampel dan Tipe Kernel.....	33



*Mencerdaskan dan
Memartabatkan Bangsa*

DAFTAR GAMBAR

Gambar 2.1. Ilustrasi Batas Keputusan yang Mungkin untuk Set Data.....	11
Gambar 2.2. Ilustrasi Margin <i>Hyperplane</i>	12
Gambar 2.3. Ilustrasi <i>Hyperplane</i> pada SVM.....	12
Gambar 2.4. Diagram Plot Data Linear	13
Gambar 2.5. Ilustrasi Dimensi Data.....	18
Gambar 2.6. Diagram Plot Non-Linear dengan Kernel Gaussian RBF.....	19
Gambar 2.7. Diagram Plot Non-Linear dengan Kernel Polinomial.....	20
Gambar 2.8. Diagram Alir Jawaban Masalah Penelitian.....	25
Gambar 3.1. Diagram Alir Penelitian	27

*Mencerdaskan dan
Memartabatkan Bangsa*