



**OPTIMASI ALGORITMA *SUPPORT VECTOR MACHINE*
MENGUNAKAN SELEKSI FITUR *PARTICLE SWARM*
OPTIMIZATION PADA ANALISIS SENTIMEN TERHADAP
KEBIJAKAN PPKM**

SKRIPSI

HASAN MUBAROK

1810511013

**UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN
JAKARTA**

**FAKULTAS ILMU KOMPUTER
PROGAM STUDI S1 INFORMATIKA**

2022



**OPTIMASI ALGORITMA *SUPPORT VECTOR MACHINE*
MENGUNAKAN SELEKSI FITUR *PARTICLE SWARM*
OPTIMIZATION PADA ANALISIS SENTIMEN TERHADAP
KEBIJAKAN PPKM**

SKRIPSI

**Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh Gelar
Sarjana Komputer**

HASAN MUBAROK

1810511013

**UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN
JAKARTA**

FAKULTAS ILMU KOMPUTER

PROGAM STUDI S1 INFORMATIKA

2022

PERNYATAAN ORISINALITAS

Tugas Akhir ini adalah hasil karya sendiri dan semua sumber yang dikutip maupun yang dirujuk telah saya nyatakan dengan benar.

Nama : Hasan Mubarak

NRP : 1810511013

Tanggal : 7 Juni 2022

Bilamana dikemudian hari ditemukan ketidaksesuaian dengan pernyataan saya ini, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan yang berlaku

Jakarta, 7 Juni 2022

Yang menyatakan



(Hasan Mubarak)

PERNYATAAN PUBLIKASI

Sebagai civitas akademik Universitas Pembangunan Nasional Veteran Jakarta, saya yang bertanda tangan dibawah ini:

Nama : Hasan Mubarok
NIM : 1810511013
Fakultas : Ilmu Komputer
Program Studi : S1-Informatika

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Pembangunan Nasional Veteran Jakarta Hak Bebas Royalti Non eksklusif (Non-Exchange Royalty Free Right) atas karya ilmiah saya yang berjudul:

**OPTIMASI ALGORITMA *SUPPORT VECTOR MACHINE*
MENGUNAKAN SELEKSI FITUR *PARTICLE SWARM OPTIMIZATION*
PADA ANALISIS SENTIMEN TERHADAP KEBIJAKAN PPKM**

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti di Universitas Pembangunan Nasional Veteran Jakarta berhak menyimpan, mengalih media/formalitas, mengelola dalam bentuk pengkalan data (Basis Data), merawat dan mempublikasi Skripsi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta, Demikian Pernyataan ini saya buat dengan sebenarnya.

Dibuat : Jakarta

Pada Tanggal : 7 Juni 2022

Yang menyatakan,



(Hasan Mubarok)

LEMBAR PENGESAHAN

Dengan ini dinyatakan bahwa Skripsi berikut:

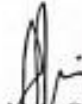
Nama : Hasan Mubarak
NIM : 1810511013
Program Studi : S1 Informatika
Judul Skripsi : Optimasi Algoritma *Support Vector Machine* Menggunakan
Seleksi Fitur *Particle Swarm Optimization* Pada Analisis
Sentimen Terhadap Kebijakan PPKM.

Telah berhasil dipertahankan di hadapan tim penguji dan diterima sebagai persyaratan yang diperlukan untuk memperoleh gelar Sarjana Komputer pada Program Studi S1 Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta.


Jayanta, S.Kom., M.Si
Ketua Penguji


Ati Zaidiah, S.Kom., MTI
Anggota Penguji


Iin Ernawati, S.Kom., M.Si
Pembimbing 1


Nurul Chamidah, S.Kom., M.Kom.
Pembimbing 2



Dr. Ermatita, M.Kom.
Dekan


Desta Sandya Prasvita, S.Kom., M.Kom
Ketua Program Studi

Diteptakan di : Jakarta
Tanggal Persetujuan : 7 Juni 2022



OPTIMASI ALGORITMA *SUPPORT VECTOR MACHINE* MENGUNAKAN SELEKSI FITUR *PARTICLE SWARM* *OPTIMIZATION* PADA ANALISIS SENTIMEN TERHADAP KEBIJAKAN PPKM

Hasan Mubarak

ABSTRAK

Twitter merupakan media sosial yang bersifat *micro-blogging* yang memungkinkan penggunaanya untuk mengekspresikan opini tentang berbagai topik dan membahas permasalahan yang terjadi saat ini. Salah satu topik yang sering diperbincangkan masyarakat yaitu penerapan kebijakan PPKM di Indonesia yang menimbulkan prokontra sehingga opini dari masyarakat sangat beragam terutama pengguna twitter. Banyaknya opini maka perlu adanya analisis sentimen. Tujuannya adalah untuk mengetahui opini masyarakat terhadap penerapan PPKM melalui tagar #PPKM. Oleh karena itu, penelitian ini dilakukan proses klasifikasi pada penerapan PPKM menggunakan dua kelas yaitu kelas sentimen positif dan kelas sentimen negatif. Metode yang digunakan dalam mengklasifikasi yakni algoritma *Support Vector Machine* dan algoritma *Particle Swarm Optimization* sebagai seleksi fitur. Kedua algoritma tersebut akan dibagi menjadi dua proses yakni membuat model menggunakan PSO dan tidak menggunakan PSO. Pengambilan dengan teknik *crawling* dimulai dari tanggal 1 Juli – 30 Agustus 2021 dengan API yang sudah disediakan oleh twitter. Hasil evaluasi klasifikasi menggunakan *confusion matrix* didapatkan nilai akurasi 79.77%, recall 69.04%, dan 85.29% pada data tanpa PSO (Seleksi Fitur). Sedangkan pada data menggunakan PSO (Seleksi Fitur) didapatkan nilai akurasi sebesar 87.08%, recall 76.83%, dan Precision 94.03%.

Kata kunci: Twitter, PPKM, *Support Vector Machine* (SVM), *Particle Swarm Optimization* (PSO).

OPTIMIZATION OF SUPPORT VECTOR MACHINE ALGORITHM USING FEATURE SELECTION OF PARTICLE SWARM OPTIMIZATION IN SENTIMENT ANALYSIS OF PPKM POLICY

Hasan Mubarak

ABSTRACT

Twitter is a micro-blogging social media that allows users to express opinions on various topics and discuss current issues. One of the topics that are often discussed by the public is the implementation of PPKM in Indonesia which raises pros and cons so that opinions from the public are very diverse, especially Twitter users. With the number of opinions, it is necessary to have a sentiment analysis. The aim is to find out public opinion on the implementation of PPKM through the hashtag #PPKM. Therefore, this research carried out the classification process on the application of PPKM using two classes, namely the positive sentiment class and the negative sentiment class. The method used in classifying is the Support Vector Machine algorithm and the Particle Swarm Optimization algorithm as feature selection. The two algorithms will be divided into two processes, namely making models using PSO and not using PSO. Retrieval by crawling technique starts from July 1 – August 30, 2021, with the API that has been provided by Twitter. The results of the classification evaluation using the confusion matrix obtained an accuracy value of 79.77%, recall 69.04%, and 85.29% on data without PSO (Feature Selection). While the data using PSO (Feature Selection) obtained an accuracy value of 87.08%, recall 76.83% and Precision 94.03%.

Kata kunci: Twitter, PPKM, *Support Vector Machine* (SVM), *Particle Swarm Optimization* (PSO).

KATA PENGANTAR

Puji serta syukur penulis panjatkan kehadiran Allah SWT atas segala limpahan rahmat, karunia dan ridha-Nya, sehingga Skripsi dengan judul “Optimasi Algoritma *Support Vector Machine* Menggunakan Seleksi Fitur *Particle Swarm Optimization* Pada Analisis Sentimen Terhadap Kebijakan PPKM” berhasil diselesaikan. Tak lupa, penulis ingin mengucapkan banyak terima kasih kepada:

1. Kedua orang tua penulis, Makmur (bapak) dan Suprapti (ibu), serta keluarga yang selalu memberikan doa dan motivasi sehingga penulis dapat menyelesaikan skripsi ini.
2. Iin Ernawati, S.Kom., M.Si. dan Nurul Chamidah, S.Kom., M.Kom. selaku dosen pembimbing yang telah membimbing serta memberikan berbagai arahan dan saran yang bermanfaat.
3. Ibu Dr. Ermatita, M. Kom. selaku Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jakarta.
4. Bapak Desta Sandya Prasvita, S.Kom., M.Kom. selaku Ketua Program Studi Informatika Universitas Pembangunan Nasional Veteran Jakarta.
5. Bapak Bapak Dr. Didit Widiyanto, S.Kom. selaku dosen pembimbing akademik.
6. Bapak/Ibu dosen Informatika Universitas Pembangunan Nasional Veteran Jakarta yang telah memberikan ilmu yang bermanfaat.
7. Seluruh teman-teman penulis yang tidak dapat disebutkan satu-persatu yang berkontribusi dalam mendukung, memberi bantuan, dan doa bagi kelancaran penelitain ini.

Akhir kata penulis ucapkan terima kasih dan semoga skripsi ini dapat bermanfaat bagi pembaca.

Jakarta

Penulis

DAFTAR ISI

OPTIMASI ALGORITMA <i>SUPPORT VECTOR MACHINE</i> MENGGUNAKAN SELEKSI FITUR <i>PARTICLE SWARM OPTIMIZATION</i> PADA ANALISIS SENTIMEN TERHADAP KEBIJAKAN PPKM.....	i
PERNYATAAN ORISINALITAS	ii
PERNYATAAN PUBLIKASI.....	iii
LEMBAR PENGESAHAN	iv
ABSTRAK	v
ABSTRACT	vi
KATA PENGANTAR	vii
DAFTAR ISI.....	viii
DAFTAR GAMBAR	xi
DAFTAR TABEL.....	xii
DAFTAR SIMBOL	xiii
 BAB I PENDAHULUAN.....	 1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan Penelitian.....	3
1.4 Ruang Lingkup Penelitian.....	3
1.5 Luaran yang Diharapkan	4
1.6 Manfaat Penelitian.....	4
1.7 Sistematika Penulisan.....	4
 BAB II TINJAUAN PUSTAKA	 7
2.1 Twitter	7
2.2 Pemberlakuan Pembatasan Kegiatan Masyarakat (PPKM)	7
2.3 Data Mining.....	7
2.4 <i>Text Mining</i>	7
2.5 Analisis Sentimen.....	8
2.6 Pra-proses Data	8

2.6.1	<i>Case Folding</i>	8
2.6.2	<i>Cleaning</i>	8
2.6.3	<i>Normalization</i>	9
2.6.4	<i>Tokenization</i>	9
2.6.5	<i>Stopword Removal</i>	9
2.6.6	<i>Stemming</i>	9
2.7	<i>Term Frequency–Inverse Document Frequency (TF-IDF)</i>	9
2.8	<i>Particle Swarm Optimization</i>	10
2.9	Support Vector Machine (SVM)	11
2.10	Evaluasi	13
2.11	Penelitian Terkait	14
BAB III METODELOGI PENELITIAN		16
3.1	Tahapan Penelitian	16
3.2	Identifikasi Masalah	17
3.3	Studi Pustaka	17
3.4	Pengumpulan Data	17
3.5	Pelabelan Data	17
3.6	Pra-proses Data	19
3.6.1	<i>Case Folding</i>	20
3.6.2	<i>Cleaning</i>	21
3.6.3	<i>Normalization</i>	21
3.6.4	<i>Tokenization</i>	22
3.6.5	<i>Stopword Removal</i>	23
3.6.6	<i>Stemming</i>	23
3.7	Pembobotan TF-IDF	24
3.8	<i>Particle Swarm Optimization</i>	24
3.9	Klasifikasi Data	25
3.10	Evaluasi	25
3.11	Perangkat Penelitian	25
3.12	Jadwal Penelitian	26

BAB IV HASIL DAN PEMBAHASAN	28
4.1 Data	28
4.2 Pelabelan Data.....	29
4.3 Pra-Proses Data	34
4.3.1 <i>Case Folding</i>	34
4.3.2 <i>Cleaning</i>	35
4.3.3 <i>Tokenization</i>	36
4.3.4 <i>Normalization</i>	37
4.3.5 <i>Stopword Removal</i>	38
4.3.6 <i>Stemming</i>	39
4.4 Pembobotan TF-IDF	39
4.5 <i>Particle Swarm Optimization</i>	42
4.6 Proses Klasifikasi	43
4.6.1 Pembagian Data	44
4.6.2 Pemodelan Support Vector Machine Tanpa PSO	45
4.6.3 Pemodelan Support Vector Machine dan PSO	45
4.7 Evaluasi	45
4.7.1 Evaluasi SVM	45
4.7.2 Evaluasi SVM + PSO.....	46
4.7.3 Hasil Perbandingan	49
4.8 Visualisasi	50
4.8.1 <i>Wordcloud</i> Sentimen Positif	51
4.8.2 <i>Wordcloud</i> Sentimen Negatif.....	51
 BAB V PENUTUP	 53
5.1 Kesimpulan.....	53
5.2 Saran.....	54
DAFTAR PUSTAKA	55
RIWAYAT HIDUP	58
LAMPIRAN.....	59

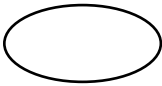
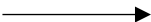
DAFTAR GAMBAR

Gambar 2. 1 Pencarian Hyperplane	11
Gambar 3. 1 Tahapan Penelitian	16
Gambar 3. 2 Skala Nilai Kappa	19
Gambar 3. 3 Pra-Proses Data	20
Gambar 4. 1 Data Hasil Crawling	28
Gambar 4. 2 Wordcloud Sentimen Positif	51
Gambar 4. 3 Wordcloud Sentimen Negatif.....	52

DAFTAR TABEL

Tabel 2. 1 Tabel Confusion Matrix.....	13
Tabel 3. 1 Contoh Case Folding	20
Tabel 3. 2 Contoh Cleaning	21
Tabel 3. 3 Contoh Normalization.....	22
Tabel 3. 4 Contoh Tokenization.....	22
Tabel 3. 5 Contoh Stopword Removal.....	23
Tabel 3. 6 Contoh Stemming	23
Tabel 3. 7 Jadwal Kegiatan	26
Tabel 4. 1 Contoh Pelabelan Manual	29
Tabel 4. 2 Perhitungan Frekuensi Positif dan Negatif	30
Tabel 4. 3 Data Sampel Pra Proses	34
Tabel 4. 4 Data Hasil Tahapan Case Folding.....	35
Tabel 4. 5 Data Hasil Tahapan Cleaning	36
Tabel 4. 6 Data Hasil Tahapan Tokenization.....	36
Tabel 4. 7 Data Hasil Tahapan Normalization.....	37
Tabel 4. 8 Data Hasil Tahapan Stopword Removal.....	38
Tabel 4. 9 Data Hasil Tahapan Stemming	39
Tabel 4. 10 Data Sampel Dokumen	40
Tabel 4. 11 Perhitungan TF – IDF	40
Tabel 4. 12 Jumlah Fitur	43
Tabel 4. 13 Pembagian Data SVM dan SVM + PSO	44
Tabel 4. 14 Confusion Matrix SVM	46
Tabel 4. 15 Confusion Matrix Iterasi 50	47
Tabel 4. 16 Confusion Matrix Iterasi 100	47
Tabel 4. 17 Confusion Matrix Iterasi 150	48
Tabel 4. 18 Confusion Matrix Iterasi 900	48
Tabel 4. 19 Confusion Matrix Iterasi 950	49
Tabel 4. 20 Confusion Matrix Iterasi 1000	49
Tabel 4. 21 Perbandingan Hasil Evaluasi	50

DAFTAR SIMBOL

Simbol <i>Flowchart</i>			
No	Simbol	Nama	Keterangan
1		<i>Terminals (Start, End)</i>	Simbol ini digunakan untuk menandakan awal dan akhir pada suatu proses.
2		<i>Processing</i>	Simbol yang menggambarkan deskripsi terhadap proses yang dijalankan.
3		<i>Flow Line</i>	Simbol yang menghubungkan antar simbol dan menyatakan jalannya proses dalam sistem.