

**PENERAPAN PRACTICAL SWARM OPTIMIZATION SVM
DENGAN TEKNIK SMOTE UNTUK KLASIFIKASI
SENTIMEN PROGRAM KERJA BUMN**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi S1 Informatika



disusun oleh

NOOR HAFIDZ AL HAITSAM I

20.11.3541

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2024

**PENERAPAN PRACTICAL SWARM OPTIMIZATION SVM
DENGAN TEKNIK SMOTE UNTUK KLASIFIKASI
SENTIMEN PROGRAM KERJA BUMN**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi S1 Informatika



disusun oleh

NOOR HAFIDZ AL HAITSAMI

20.11.3541

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2024

HALAMAN PERSETUJUAN

SKRIPSI

**PENERAPAN PRACTICAL SWARM OPTIMIZATION SVM
DENGAN TEKNIK SMOTE UNTUK KLASIFIKASI
SENTIMEN PROGRAM KERJA BUMN**


yang disusun dan diajukan oleh

NOOR HAFIDZ AL-HAITSAMI

20.11.3541

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 12 Agustus 2024

Dosen Pembimbing,


Akhmad Dahlan, M.Kom
NIK. 190302174

HALAMAN PENGESAHAN

SKRIPSI

PENERAPAN PRACTICAL SWARM OPTIMIZATION SVM
DENGAN TEKNIK SMOTE UNTUK KLASIFIKASI
SENTIMEN PROGRAM KERJA BUMN

yang disusun dan diajukan oleh

NOOR HAFIDZ AL-HAITSAMI

20.11.3541

Telah dipertahankan di depan Dewan Penguji
pada tanggal 12 Agustus 2024

Susunan Dewan Penguji

Nama Penguji

Majid Rahardi, S.Kom., M.Eng
NIK. 190302393

Dr. Fery Wahyu Wibowo, S.Si., M.Cs.
NIK. 190302235

Ahmad Dahlan, M.Kom
NIK. 190302174

Tanda/Tangan

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 12 Agustus 2024

DEKAN FAKULTAS ILMU KOMPUTER



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

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : NOOR HAFIDZ AL-HAITSAMI
NIM : 20.11.3541

Menyatakan bahwa Skripsi dengan judul berikut:

**PENERAPAN PRACTICAL SWARM OPTIMIZATION SYM DENGAN
TEKNIK SMOTE UNTUK KLASIFIKASI SENTIMEN
PROGRAM KERJA BUMN**

Dosen Pembimbing : Akhmad Dahlan, M.Kom

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, 12 Agustus 2024

Yang Menandatangani,



NOOR Hafidz Al-Haitsami

HALAMAN PERSEMBAHAN

Pertama saya ucapkan puji syukur kepada Allah SWT atas segala nikmat dan kasih sayangnya. Skripsi ini penulis persembahkan untuk:

1. Papi dan Mami yang sangat saya sayangi, di mana telah memberikan kasih sayang yang berlimpah terhadap penulis sehingga bisa membuat penulis bertahan hidup sampai sekarang. Terimakasih telah menjadi orangtua yang sangat hebat.
2. Akhmad Dahlan, M.Kom. selaku dosen pembimbing saya yang sangat baik dan mau membimbing saya selama ini. Terima kasih atas setiap waktu, tenaga, dan perhatian yang Bapak berikan. Semoga segala kebaikan Bapak mendapatkan balasan yang setimpal dari Tuhan Yang Maha Esa.
3. Untuk Ghaluh, Seli, Raniya, Ardhiansa, Rangga, Dimas alias teman di bangku kuliah yang telah menemani *ups and downs* dari hidup penulis, terimakasih telah mendengarkan keluh-kesah serta mengerti dan memahami penulis setiap saat.
4. *Last but not least, I wanna thank me, I wanna thank me for believing in me, I wanna thank me for doing all this hard work, I wanna thank me for having no days off, I wanna thank me for, for never quitting. I wanna thank me all the time.*

KATA PENGANTAR

Segala puji bagi Allah SWT atas limpahan rahmat, petunjuk dan karunia-Nya yang telah membantu peneliti menyelesaikan penyusunan skripsi ini. Skripsi ini disusun sebagai tahapan penting untuk memenuhi persyaratan kelulusan program Studi Informatika Fakultas Ilmu Komputer Universitas Amikom Yogyakarta. Penyusunan skripsi ini tidak lepas dari doa, bantuan, dukungan dan bimbingan dari berbagai pihak. Oleh karena itu, penulis ingin mengucapkan terimakasih yang sebesar-besarnya kepada:

1. Keluarga tersayang, yang senantiasa memberikan doa terbaik, dukungan moral, materi, serta semangat dalam perjalanan penyusunan skripsi ini.
2. Akhmad Dahlan, M.Kom, selaku dosen pembimbing, atas dukungan, arahan, serta masukan berharga yang telah membimbing langkah-langkah penyusunan skripsi ini.
3. Sahabat dan teman, yang telah berbagi ilmu, pengalaman, dan inspirasi sepanjang perjalanan akademik di Universitas Amikom Yogyakarta.
4. Semua individu yang turut menyumbangkan gagasan, pandangan, dan sokongan dalam berbagai bentuk, yang telah membantu kelancaran perjalanan penyelesaian skripsi ini.

Yogyakarta, 12 agustus 2024

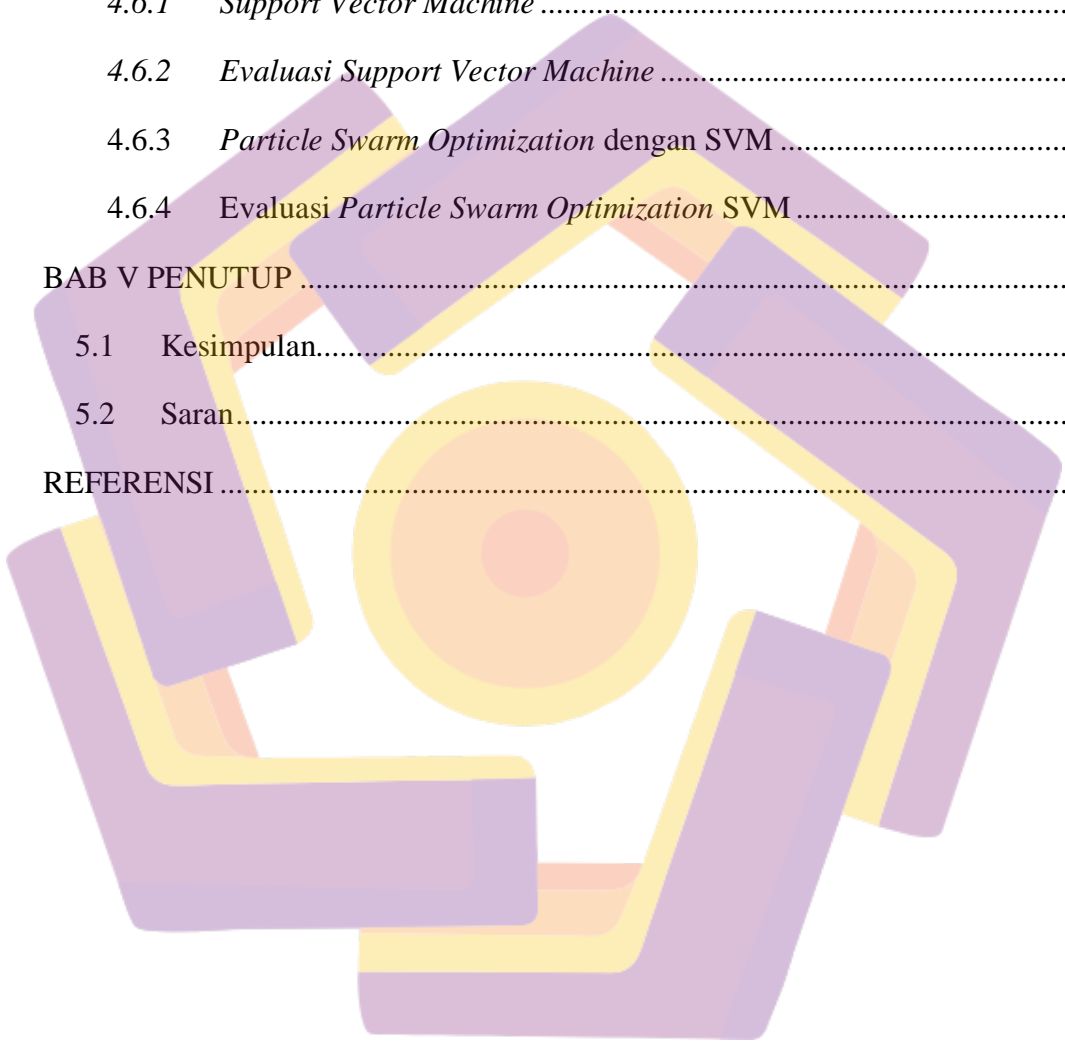
Penulis

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI.....	vii
DAFTAR TABEL.....	x
DAFTAR GAMBAR	xi
DAFTAR LAMPIRAN.....	xii
DAFTAR LAMBANG DAN SINGKATAN	xii
DAFTAR ISTILAH	xiv
INTISARI	xv
ABSTRACT.....	xvi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	3
1.3 Batasan Masalah.....	3
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	4
1.6 Sistematika Penulisan.....	5
BAB II TINJAUAN PUSTAKA	6
2.1 Studi Literatur	6

2.2	Analisis Sentiment.....	15
2.3	Text Mining.....	15
2.4	Instagram.....	15
2.5	Badan Usaha Milik Negara	16
2.6	Pra Proses Teks	16
2.7	<i>Synthetic Minority Over-sampling Technique</i>	17
2.8	TF-IDF	17
2.9	Support Vector Machine	18
2.10	<i>Particle Swarm Optimization (PSO)</i>	18
2.11	Confussion Matriks.....	19
BAB III METODE PENELITIAN		21
3.1	Objek Penelitian.....	21
3.2	Alur Penelitian.....	21
3.2.1	Pengumpulan Dataset.....	22
3.2.2	Pra-proses Teks.....	24
3.2.3	Labeling	26
3.2.4	TF-IDF	26
3.2.5	<i>Synthetic Minority Over-sampling Technique</i>	26
3.2.6	Pembagian Data	27
3.2.7	Klasifikasi	27
3.2.8	Confussion Matriks.....	28
3.2.9	Laporan Penelitian	28
BAB IV HASIL DAN PEMBAHASAN		30
4.1	Teknik Pengumpulan Data.....	30
4.2	Pra Proses Text.....	32

4.3	Pelabelan	36
4.4	Vektorisasi TF-IDF	39
4.5	Penerapan SMOTE (<i>Synthetic Minority Over-sampling Technique</i>).....	39
4.6	Modeling	40
4.6.1	<i>Support Vector Machine</i>	41
4.6.2	<i>Evaluasi Support Vector Machine</i>	43
4.6.3	<i>Particle Swarm Optimization</i> dengan SVM	45
4.6.4	<i>Evaluasi Particle Swarm Optimization SVM</i>	45
BAB V PENUTUP		54
5.1	Kesimpulan.....	54
5.2	Saran.....	55
REFERENSI		56



DAFTAR TABEL

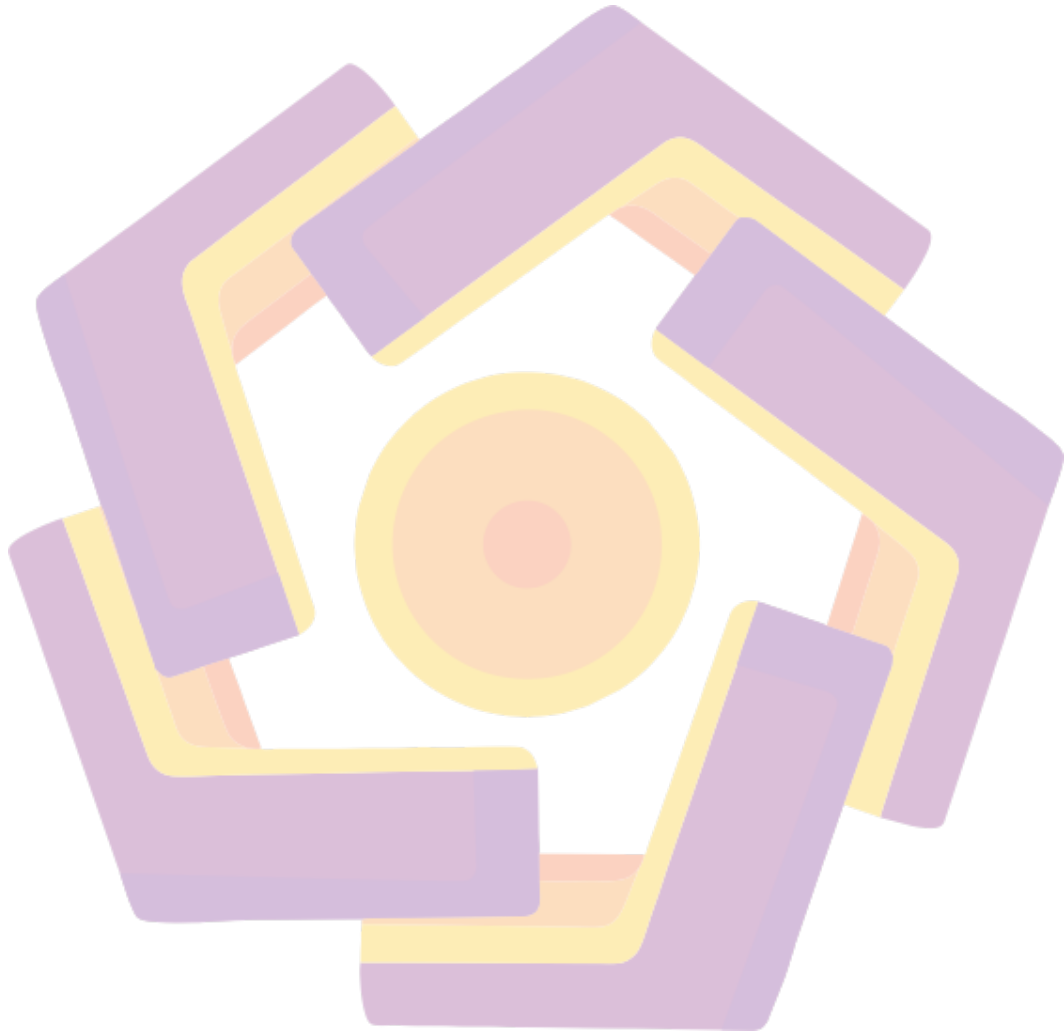
Tabel 2. 1 Keaslian Penelitian	10
Tabel 2. 2 Confussion Matriks.....	20
Tabel 3. 1 Corpus Normalisasi Kata.....	23
Tabel 4. 1 Sampel Data Hasil Crawling.....	31
Tabel 4. 2 Hasil Case Folding.....	32
Tabel 4. 3 Hasil Cleaning.....	33
Tabel 4. 4 Hasil Normalisasi Kata.....	33
Tabel 4. 5 Hasil Stemming.....	34
Tabel 4. 6 Hasil Tokenizing.....	35
Tabel 4. 7 Hasil Stopword Removal.....	35
Tabel 4. 8 Hasil Penilaian Sentimen.....	37
Tabel 4. 9 Parameter SVM.....	41
Tabel 4. 10 Hasil Grid Search Parameter SVM.....	41
Tabel 4. 11 Hasil Evaluasi SVM.....	43
Tabel 4. 12 Confussion Matrik SVM.....	43
Tabel 4. 13 Fitness Test_size 0.3.....	46
Tabel 4. 14 Fitness Test_size 0.2.....	47
Tabel 4. 15 Fitness Test_size 0.1.....	48
Tabel 4. 16 Evaluasi PSO SVM.....	49

DAFTAR GAMBAR

Gambar 1. 1 Data Pengguna Instagram Setiap Negara.....	1
Gambar 2. 1 <i>Hyperplane Supervector Machine</i>	18
Gambar 3. 1 Alur Penelitian	22
Gambar 4. 1 <i>IGCommentsExport - Export Comment for IG.</i>	30
Gambar 4. 2 Distribusi Poliritas.....	38
Gambar 4. 3 World Cloud Positif dan Negatif	38
Gambar 4. 4 Vektorisasi TF-IDF	39
Gambar 4. 5 Penerapan Teknik Smote	40
Gambar 4. 6 Fungsi Optimasi PSO.....	45
Gambar 4. 7 Parameter PSO	46
Gambar 4. 8 Grafik Perbandingan SVM dan PSO SVM.....	52

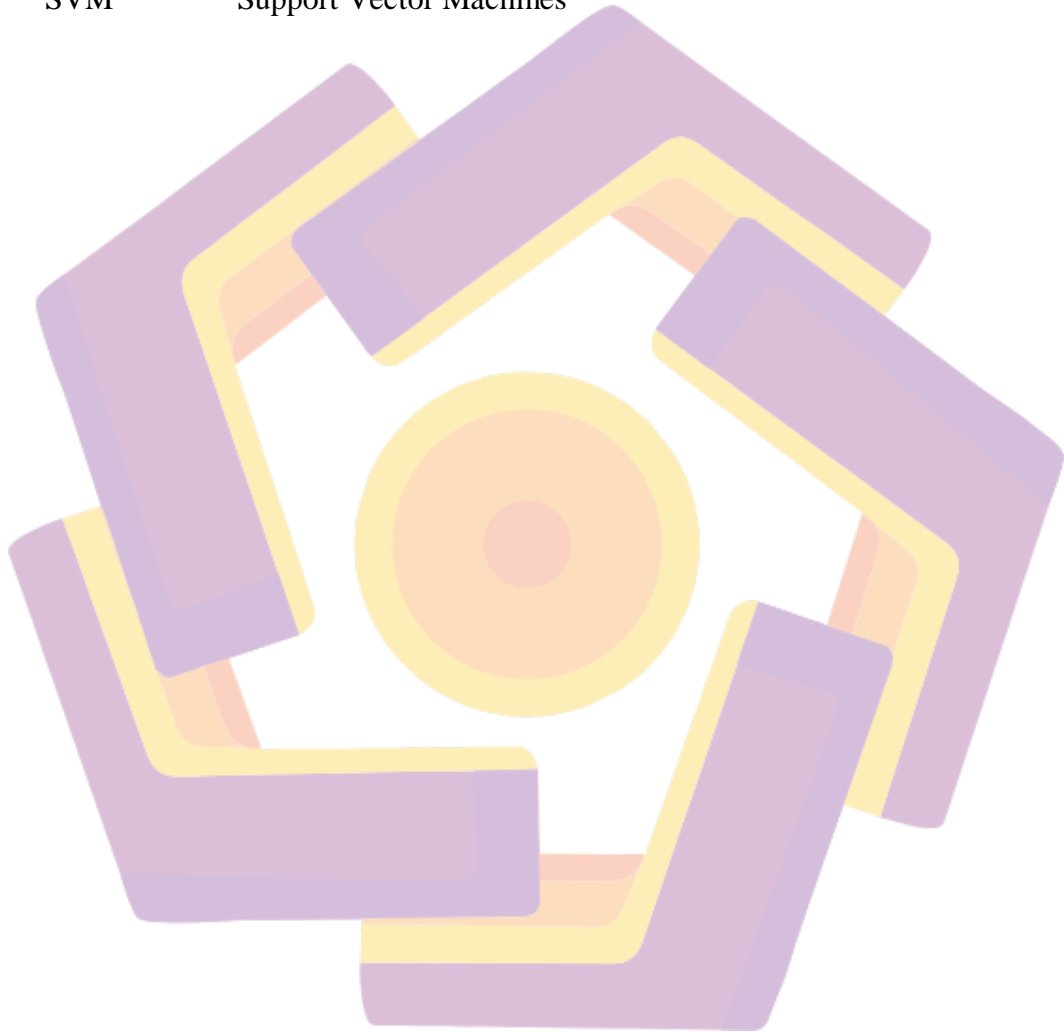
DAFTAR LAMPIRAN

Lampiran 1. Profil obyek Penelitian	10
Lampiran 2. Dokumentasi Penelitian	11



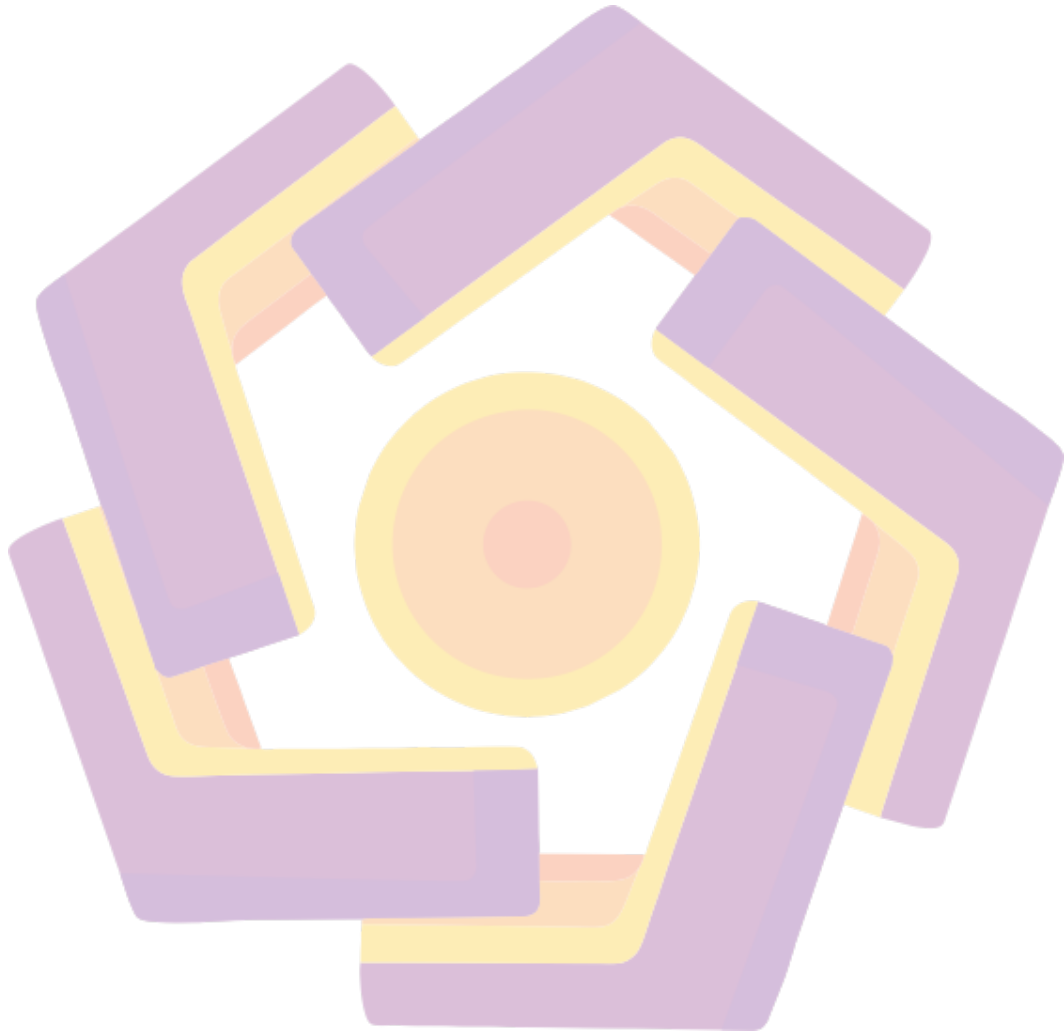
DAFTAR LAMBANG DAN SINGKATAN

Ω	Tahanan Listrik
μ	Konstanta gesekan
ANFIS	Adaptive Network Fuzzy Inference System
SVM	Support Vector Machines



DAFTAR ISTILAH

Vektor	besaran yang mempunyai arah
Eigen Value	akar akar persamaan



INTISARI

Penelitian ini bertujuan untuk mengembangkan model klasifikasi sentimen terhadap program kerja BUMN dengan menggunakan algoritma Support Vector Machine (SVM) yang dioptimalkan oleh Particle Swarm Optimization (PSO) dan teknik SMOTE (Synthetic Minority Over-sampling Technique). Data dikumpulkan dari dua akun Instagram resmi, yaitu FHCI BUMN dan KEMENTERIAN BUMN, dengan total 3000 komentar. Proses pre-processing mencakup case folding, cleaning, normalisasi, stemming, tokenizing, dan stopword removal. Pelabelan dilakukan dengan metode manual. Hasilnya, ditemukan bahwa 1686 komentar bersentimen positif dan 1048 bersentimen negatif. Teknik SMOTE kemudian diterapkan untuk menyeimbangkan distribusi kelas, menghasilkan data yang seimbang dan siap untuk dianalisis lebih lanjut. Sebelum penerapan PSO pada SVM, model SVM biasa menghasilkan akurasi tertinggi sebesar 75% dengan recall dan precision masing-masing sebesar 76% dan 75% pada skema pembagian data 90:10. Model SVM dioptimalkan menggunakan PSO untuk menemukan parameter optimal, yaitu C dan gamma, dengan hasil evaluasi menunjukkan akurasi tertinggi sebesar 82% pada skema pembagian data 90:10. Hasil ini didukung dengan nilai precision dan recall yang baik untuk kedua kelas sentimen, dengan recall mencapai 82% dan precision mencapai 82%. Penelitian ini menegaskan bahwa PSO efektif untuk mengoptimalkan parameter pada SVM, menghasilkan model klasifikasi sentimen yang lebih akurat dan andal. Pendekatan ini berpotensi untuk diterapkan dalam berbagai aplikasi yang memerlukan prediksi sentimen yang presisi dan dapat diandalkan.

Kata kunci: analisis sentiment, smote, support vector machine, pso

ABSTRACT

This study aims to develop a sentiment classification model for BUMN (State-Owned Enterprises) work programs using a Support Vector Machine (SVM) algorithm optimized by Particle Swarm Optimization (PSO) and the Synthetic Minority Over-sampling Technique (SMOTE). Data were collected from two official Instagram accounts, namely FHCI BUMN and KEMENTERIAN BUMN, totaling 3000 comments. The pre-processing process includes case folding, cleaning, normalization, stemming, tokenizing, and stopword removal. Labeling was performed manually. The results show that 1686 comments were classified as positive sentiment and 1048 as negative sentiment. The SMOTE technique was then applied to balance the class distribution, resulting in a balanced dataset ready for further analysis. Before applying PSO to SVM, the standard SVM model achieved a highest accuracy of 75% with recall and precision of 76% and 75%, respectively, in the 90:10 data split scheme. The SVM model was optimized using PSO to find the optimal parameters, C and gamma, with the evaluation results showing the highest accuracy of 82% in the 90:10 data split scheme. This was supported by good precision and recall values for both sentiment classes, with recall reaching 82% and precision also reaching 82%. This study confirms that PSO is effective for optimizing parameters in SVM, resulting in a more accurate and reliable sentiment classification model. This approach has potential applications in various scenarios requiring precise and reliable sentiment prediction.

Keyword: *sentiment analysis, smote, support vector machine, pso*