

## DAFTAR ISI

HALAMAN PENGESAHAN.....	i
PERNYATAAN.....	ii
HALAMAN MOTIVASI.....	iii
PRAKATA.....	iv
DAFTAR ISI.....	vi
DAFTAR TABEL.....	ix
DAFTAR GAMBAR.....	x
INTISARI.....	xi
ABSTRACT.....	xii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	4
1.3 Batasan Masalah.....	4
1.4 Tujuan Penelitian.....	5
1.5 Manfaat Penelitian.....	5
1.6 Metode Penelitian.....	5
1.7 Sistematika Penulisan.....	6
BAB II TINJAUAN PUSTAKA.....	6
BAB III LANDASAN TEORI.....	10
3.1 SMS Spam.....	10
3.2 Klasifikasi.....	11
3.3 <i>Preprocessing</i> .....	12
3.3.1 <i>Case Folding</i> .....	12
3.3.2 Tokenisasi.....	13
3.3.3 <i>Filtering</i> .....	13
3.3.4 <i>Stemming</i> .....	13
3.4 Ekstraksi dan Seleksi Fitur.....	14
3.4.1 <i>Document Term Matrix</i> .....	14
3.5 Particle Swarm Optimization.....	15
3.5.1 Fungsi Objektif.....	16
3.5.2 <i>Velocity</i> .....	16
3.5.3 <i>Position</i> .....	17
3.5.4 <i>Binary PSO</i> .....	18

3.6	Naive Bayes .....	18
3.6.1	Gaussian Naive Bayes.....	19
3.7	K-Fold Cross Validation .....	19
3.8	Parameter Evaluasi.....	20
3.8.1	Akurasi .....	20
3.8.2	Presisi .....	21
3.8.3	Recall.....	21
3.8.4	F1 Score .....	21
BAB IV Analisis dan rancangan .....		22
4.1	Deskripsi Umum Penelitian .....	22
4.2	Tahapan Penelitian .....	23
4.3	Pengumpulan Data .....	24
4.4	Rancangan Model.....	24
4.4.1	Preprocessing .....	25
4.4.2	Model PSO .....	28
4.4.3	Model Gaussian Naïve Bayes .....	30
4.5	Pengujian.....	31
BAB V IMPLEMENTASI.....		33
5.1	Implementasi .....	33
5.2	Pengumpulan Data dan Pelabelan Manual.....	34
5.2.1	Implementasi Preprocessing .....	34
5.3	Klasifikasi dengan Gaussian Naive Bayes.....	36
5.4	Implementasi Seleksi Fitur dengan <i>Particle Swarm Optimization</i> .....	39
BAB VI HASIL DAN PEMBAHASAN .....		44
6.1	Hasil Pelabelan.....	44
6.2	Hasil Preprocessing.....	45
6.3	Hasil Gaussian Naive Bayes .....	47
6.4	Hasil Seleksi Fitur dengan PSO .....	50
6.5	Hasil validasi dan evaluasi setelah PSO .....	54
BAB VII KESIMPULAN DAN SARAN .....		61
7.1	Kesimpulan .....	61
7.2	Saran.....	61
DAFTAR PUSTAKA .....		63



UNIVERSITAS  
GADJAH MADA

**DETEKSI SPAM PADA SMS BAHASA INDONESIA MENGGUNAKAN SELEKSI FITUR BERBASIS  
PARTICLE SWARM  
OPTIMIZATION**

Dyah Mawar Umbulsari, Anny Kartika Sari, S.Si., M.Kom.,Ph.D

Universitas Gadjah Mada, 2020 | Diunduh dari <http://etd.repository.ugm.ac.id/>

LAMPIRAN .....67

## DAFTAR TABEL

Tabel 1 Perbandingan penelitian-penelitian sebelumnya.....	9
Tabel 2 Contoh SMS SPAM.....	11
Tabel 3 Ilustrasi DTM.....	15
Tabel 4 Tabel confusion matrix .....	20
Tabel 4.1 Nilai variabel untuk menghitung <i>velocity</i> PSO.....	29
Tabel 6.1 Cuplikan hasil pelabelan data manual.....	44
Tabel 7.2 Cuplikan hasil pelabelan data publik .....	45
Tabel 6.3 Hasil akurasi untuk dataset manual.....	47
Tabel 6.3 (lanjutan) .....	48
Tabel 6.4 Hasil evaluasi untuk dataset manual .....	48
Tabel 6.4 (lanjutan) .....	49
Tabel 6.5 Hasil akurasi untuk dataset publik .....	49
Tabel 6.6 Hasil evaluasi untuk dataset publik.....	50
Tabel 6.7 Hasil akurasi setelah PSO untuk dataset manual .....	54
Tabel 6.7 (lanjutan) .....	55
Tabel 6.8 Hasil evaluasi setelah PSO untuk dataset manual.....	55
Tabel 6.8 (lanjutan) .....	56
Tabel 6.9 Hasil akurasi setelah PSO untuk dataset publik.....	56
Tabel 6.18 Hasil evaluasi setelah PSO untuk dataset publik .....	57
Tabel A.1 Data SMS yang telah dikumpulkan.....	67

## DAFTAR GAMBAR

Gambar 3.1 Ilustrasi pergerakan partikel dalam PSO (Macedo, 2018) .....	16
Gambar 3.2 Ilustrasi distribusi Gaussian .....	19
Gambar 4.1 Diagram alir proses utama model.....	25
Gambar 4.2 Ilustrasi alir preprocessing .....	26
Gambar 4.3 Ilustrasi penghapusan angka .....	26
Gambar 4.4 Ilustrasi penghapusan huruf $\leq 3$ .....	27
Gambar 4.5 Ilustrasi penghapusan elemen selain huruf .....	27
Gambar 4.6 Ilustrasi stemming .....	28
Gambar 4.7 Diagram alir <i>Binary</i> PSO .....	30
Gambar 4.8 Diagram alir Gaussian Naive Bayes.....	31
Gambar 5.1 Cuplikan implementasi preprocessing .....	35
Gambar 5.2 Implementasi stemming dan vektorisasi .....	36
Gambar 5.3 Deklarasi fungsi dan variabel untuk klasifikasi .....	37
Gambar 5.4 Fungsi utama untuk klasifikasi dan validasi .....	38
Gambar 5.5 Implementasi fungsi objektif pada PSO.....	39
Gambar 5.7 Implementasi fungsi PSO.....	40
Gambar 5.8 Implementasi fungsi utama PSO.....	40
Gambar 5.9 Implementasi klasifikasi dan validasi model baru .....	42
Gambar 5.10 Implementasi pembuatan plot untuk <i>cost</i> .....	43
Gambar 6.1 Cuplikan sebelum implementasi preprocessing dengan regex.....	46
Gambar 6.2 Cuplikan setelah implementasi preprocessing dengan regex.....	46
Gambar 6.3 Cuplikan hasil Term Frequency dan Vektorisasi .....	47
Gambar 6.4 Hasil run PSO pada dataset manual .....	51
Gambar 6.5 Grafik penurunan <i>cost</i> selama iterasi dalam PSO untuk dataset manual .....	52
Gambar 6.6 Perubahan jumlah fitur pada dataset manual.....	52
Gambar 6.7 Hasil run PSO untuk dataset publik .....	53
Gambar 6.8 Grafik penurunan <i>cost</i> selama iterasi dalam PSO untuk dataset publik.....	53
Gambar 6.9 Perubahan jumlah fitur pada dataset publik.....	54
Gambar 6.10 Grafik perbandingan akurasi & waktu sebelum dan sesudah PSO pada dataset manual .....	57
Gambar 6.11 Grafik perbandingan akurasi & waktu sebelum dan sesudah PSO pada dataset publik.....	58
Gambar 6.12 Grafik perbandingan presisi, <i>recall</i> & <i>f1-score</i> sebelum dan sesudah PSO pada dataset manual .....	59
Gambar 6.13 Grafik perbandingan presisi, <i>recall</i> & <i>f1-score</i> sebelum dan sesudah PSO pada dataset publik.....	60