



**IMPLEMENTASI *PARTICLE SWARM OPTIMIZATION* PADA
ANALISIS SENTIMEN *CYBERBULLYING* MENGGUNAKAN
*RANDOM FOREST***

Skripsi

**Untuk Memenuhi Persyaratan
Dalam Menyelesaikan Sarjana Strata-1 Ilmu Komputer**

Oleh

HELMA HERLINDA

NIM 1811016320011

**PROGRAM STUDI S-1 ILMU KOMPUTER
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMBUNG MANGKURAT
BANJARBARU
AGUSTUS 2023**



**IMPLEMENTASI *PARTICLE SWARM OPTIMIZATION* PADA
ANALISIS SENTIMEN *CYBERBULLYING* MENGGUNAKAN
*RANDOM FOREST***

Skripsi

**Untuk Memenuhi Persyaratan
Dalam Menyelesaikan Sarjana Strata-1 Ilmu Komputer**

Oleh

HELMA HERLINDA

NIM 1811016320011

**PROGRAM STUDI S-1 ILMU KOMPUTER
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMBUNG MANGKURAT
BANJARBARU
AGUSTUS 2023**

SKRIPSI

**IMPLEMENTASI *PARTICLE SWARM OPTIMIZATION* PADA ANALISIS SENTIMEN
CYBERBULLYING MENGGUNAKAN *RANDOM FOREST***

Oleh:
HELMA HERLINDA
1811016320011

Susunan Dosen Penguji:

Pembimbing I



M. Itqan Mazdadi, S.Kom., M.Kom.
NIP. 1990061220190301013

Dosen Penguji I



Dwi Kartini, S.Kom., M.Kom.
NIP. 198704212012122003

Pembimbing II



Muliadi, S.Kom., M.Sc.
NIP. 197804222010121002

Dosen Pengdji II



Irwan Budiman, S.T., M.Kom.
NIP. 197703252008121001



Surakarta, 22 Agustus 2023
Program Studi Ilmu Komputer

Budiman, S.T., M.Kom.
NIP. 197703252008121001

PERNYATAAN

Dengan ini saya menyatakan bahwa dalam skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diberikan orang lain, kecuali yang secara tertulis diacu di dalam naskah ini dan disebutkan dalam Daftar Pustaka.

Banjarbaru, Agustus 2023



HELMA HERLINDA

NIM. 1811016320011

ABSTRAK

IMPLEMENTASI PARTICLE SWARM OPTIMIZATION PADA ANALISIS SENTIMEN CYBERBULLYING MENGGUNAKAN RANDOM FOREST

(Oleh: Helma Herlinda; Pembimbing: Muhammad Itqan Mazdadi, S.Kom., M.Kom. dan Muliadi, S.Kom., M.Sc; 2023; 105 halaman)

Media sosial telah memberikan pengaruh yang signifikan pada kehidupan sebagian besar individu di era kontemporer. Ini tidak hanya memungkinkan komunikasi di antara orang-orang dalam lingkungan tertentu tetapi juga memfasilitasi konektivitas pengguna di dunia maya. Instagram adalah platform media sosial yang memainkan peran penting dalam berbagi informasi dan membina komunikasi di antara para penggunanya melalui media foto dan video, yang dapat dikomentari oleh pengguna lain. Pemanfaatan Instagram terus meningkat setiap tahunnya, sehingga berpotensi memberikan konsekuensi positif maupun negatif. Salah satu konsekuensi negatif yang sering muncul adalah *cyberbullying*. Melakukan analisis sentimen pada data *cyberbullying* dapat memberikan wawasan tentang keefektifan metodologi yang digunakan. Penelitian ini merupakan penelitian eksperimen yang bertujuan untuk membandingkan kinerja Random Forest dan Random Forest setelah menerapkan teknik seleksi fitur Particle Swarm Optimization pada tiga komposisi pemisahan data yang berbeda, yaitu 70:30, 80:20, dan 90:10. Hasil evaluasi menunjukkan bahwa skor akurasi tertinggi dicapai pada konfigurasi data split 90:10. Secara spesifik, model Random Forest menghasilkan akurasi sebesar 87,50%, sedangkan model Random Forest setelah menjalani seleksi fitur menggunakan algoritma Particle Swarm Optimization, mencapai akurasi sebesar 92,19%. Oleh karena itu, penerapan Particle Swarm Optimization sebagai teknik pemilihan fitur menunjukkan potensi untuk meningkatkan akurasi metode Random Forest.

Kata kunci: RapidMiner, Media Sosial, *Data Science*, *Text Mining*, Klasifikasi

ABSTRACT

IMPLEMENTATION OF PARTICLE SWARM OPTIMIZATION ON SENTIMENT ANALYSIS OF CYBERBULLYING USING RANDOM FOREST

*(By: Helma Herlinda; Supervisor: Muhammad Itqan Mazdadi, S.Kom., M.Kom. and
Muliadi, S.Kom., M.Sc; 2023; 105 pages)*

Social media has exerted a significant influence on the lives of the majority of individuals in the contemporary era. It not only enables communication among people within specific environments but also facilitates user connectivity in the virtual realm. Instagram is a social media platform that plays a pivotal role in the sharing of information and fostering communication among its users through the medium of photos and videos, which can be commented on by other users. The utilization of Instagram is consistently growing each year, thereby potentially yielding both positive and negative consequences. One prevalent negative consequence that frequently arises is cyberbullying. Conducting sentiment analysis on cyberbullying data can provide insights into the effectiveness of the employed methodology. This research was conducted as an experimental research, aiming to compare the performance of Random Forest and Random Forest after applying the Particle Swarm Optimization feature selection technique on three distinct data split compositions, namely 70:30, 80:20, and 90:10. The evaluation results indicate that the highest accuracy scores were achieved in the 90:10 data split configuration. Specifically, the Random Forest model yielded an accuracy of 87.50%, while the Random Forest model, after undergoing feature selection using the Particle Swarm Optimization algorithm, achieved an accuracy of 92.19%. Therefore, the implementation of Particle Swarm Optimization as a feature selection technique demonstrates the potential to enhance the accuracy of the Random Forest method.

Keywords: *RapidMiner, Social Media, Data Science, Text Mining, Classification*

PRAKATA

Assalamu'alaikum Warrahmatullahi Wabarakatuh.

Puji syukur penulis panjatkan ke hadirat Allah SWT karena atas berkat rahmat dan hidayah-Nya penulis dapat menyelesaikan skripsi yang berjudul “Penerapan Seleksi Fitur *Particle Swarm Optimization* Pada Klasifikasi Analisis Sentimen *Cyberbullying* Di Media Sosial Menggunakan Metode *Random Forest*” untuk memenuhi syarat dalam menyelesaikan pendidikan program S-1 Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lambung Mangkurat. Tak lupa pula penulis panjatkan shalawat serta salam pada junjungan nabi besar Muhammad SAW beserta para sahabat, keluarga, dan pengikut beliau hingga yaumul qiyamah.

Pada lembar ini penulis ingin mengucapkan ucapan terima kasih kepada pihak-pihak yang sangat mendukung penulis dalam pembuatan dan penyusunan skripsi ini, adapun yang dimaksud adalah sebagai berikut”

1. Keluarga terutama orang tua yang selalu memberikan bantuan, semangat, motivasi, doa, serta dukungan dalam proses penyelesaian skripsi ini.
2. Bapak Muhammad Itqan Mazdadi, S.Kom., M.Kom., selaku dosen pembimbing utama yang turut serta dalam membantu dan meluangkan waktu demi kelancaran penyelesaian skripsi ini.
3. Bapak Muliadi, S.Kom., M.Sc., selaku dosen pembimbing pendamping yang turut serta dalam membantu dan meluangkan waktu demi kelancaran penyelesaian skripsi ini.
4. Ibu Dwi Kartini, S.Kom., M.Kom., dan Bapak Rahmat Ramadhani, S.Kom., M.Sc., selaku dosen penguji yang turut membantu menguji dan memberikan saran pada penelitian saya, dan bapak Irwan Budiman, S.T., M.Kom, selaku dosen penguji pengganti.
5. Bapak Irwan Budiman S.T., M.Kom., selaku Ketua Program Studi Ilmu Komputer FMIPA ULM, dan bapak Mohammad Reza Faisal, , S.T., M.T.,

Ph.D., selaku dosen pembimbing akademik yang telah memberikan izin sehingga pengerjaan skripsi ini dapat terselesaikan.

6. Seluruh Dosen dan staf Program Studi Ilmu Komputer FMIPA ULM atas ilmu dan bantuan yang diberikan selama ini yang sangat bermanfaat.
7. Teman-teman dan sahabat-sahabat keluarga Ilmu Komputer angkatan 2018 yang selalu memberikan bantuan, semangat, dan hiburan di saat kehilangan semangat. Terima kasih untuk kenangan yang telah kalian berikan.
8. Semua pihak yang tidak dapat disebutkan satu persatu yang telah turut membantu baik secara moril maupun materil dalam penyelesaian skripsi ini.

Akhir kata penulis menyadari sepenuhnya bahwa skripsi ini masih belum sempurna, namun penulis mengharapkan bantuan berupa saran dan kritik yang membangun dari semua pihak demi kesempurnaan dan mutu penulisan skripsi ini. Semoga tulisan ini dapat bermanfaat bagi ilmu pengetahuan dan pembaca khususnya serta mendapat keridhaan Allah SWT.

Banjarbaru, Agustus 2023



Penulis

DAFTAR ISI

SKRIPSI.....	i
PERNYATAAN.....	ii
ABSTRAK.....	iii
<i>ABSTRACT</i>	iv
PRAKATA.....	v
DAFTAR ISI.....	vii
DAFTAR TABEL.....	ix
DAFTAR GAMBAR.....	x
DAFTAR LAMPIRAN.....	xi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	4
1.3 Batasan Masalah.....	4
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	5
BAB II TINJAUAN PUSTAKA.....	6
2.1 Kajian Terdahulu.....	6
2.2 Keaslian Penelitian.....	8
2.3 Instagram.....	10
2.4 Cyberbullying.....	10
2.5 Text Mining.....	12
2.6 Analisis Sentimen.....	14
2.7 Ekstraksi Fitur TF-IDF.....	15
2.8 Seleksi Fitur Particle Swarm Optimization.....	16
2.9 Random Forest.....	19
2.10 Performa Klasifikasi.....	23

BAB III METODE PENELITIAN.....	25
3.1 Alat Penelitian	25
3.2 Bahan Penelitian.....	25
3.3 Prosedur Penelitian	26
BAB IV HASIL DAN PEMBAHASAN	28
4.1 Hasil.....	28
4.1.1 Pengumpulan Data.....	28
4.1.2 <i>Pre-processing</i> Data	28
4.1.2.1 Cleaning.....	29
4.1.2.2 Case Folding	30
4.1.2.3 Tokenisasi	31
4.1.2.4 Normalisasi	32
4.1.2.5 <i>Stopword removal</i>	33
4.1.2.6 <i>Stemming</i>	35
4.1.3 Ekstraksi Fitur.....	35
4.1.4 Seleksi Fitur	36
4.1.5 Klasifikasi	37
4.1.6 Evaluasi	38
4.2 Pembahasan	39
BAB V PENUTUP.....	43
5.1 Kesimpulan	43
5.2 Saran	43
DAFTAR PUSTAKA	44
LAMPIRAN	

DAFTAR TABEL

Tabel	Halaman
Tabel 1. Keaslian Penelitian.....	8
Tabel 2. Rancangan Penelitian.....	9
Tabel 3. Model <i>confusion matrix</i> menurut (Setiawan et al., 2020).....	24
Tabel 4. Data komentar.....	28
Tabel 5. Karakter yang tidak diperlukan.....	29
Tabel 6. Hasil data <i>cleaning</i>	29
Tabel 7. Hasil data <i>case folding</i>	30
Tabel 8. Data hasil tokenisasi.....	31
Tabel 9. Kamus bahasa baku.....	32
Tabel 10. Data hasil normalisasi.....	32
Tabel 11. Kamus <i>stopword removal</i>	33
Tabel 12. Data Hasil <i>Stopword removal</i>	34
Tabel 13. Data hasil <i>stemming</i>	35
Tabel 14. Dokumen.....	36
Tabel 15. Perbandingan Hasil Akurasi.....	38
Tabel 16. Pengaruh seleksi fitur.....	42

DAFTAR GAMBAR

Gambar	Halaman
Gambar 1. Algoritma sederhana <i>Random Forest</i>	20
Gambar 2. Contoh pohon <i>Random Forest</i>	23
Gambar 3. Alur penelitian.....	26
Gambar 4. <i>Particle Swarm Optimization</i>	37
Gambar 5. Proses dalam <i>Particle Swarm Optimization</i>	37
Gambar 6. <i>Random Forest</i>	38

DAFTAR LAMPIRAN

- Lampiran 1. Kamus normalisasi
- Lampiran 2. Kamus Stopword Removal
- Lampiran 3. Sastrawi
- Lampiran 4. *Library*
- Lampiran 5. *Import Data*
- Lampiran 6. *Source Code Cleaning*
- Lampiran 7. *Source Code Case Folding*
- Lampiran 8. *Source Code* Tokenisasi
- Lampiran 9. *Source Code* Normalisasi
- Lampiran 10. *Source Code Stopword Removal*
- Lampiran 11. *Source Code Stemming*
- Lampiran 12. *Convert to CSV*
- Lampiran 13. *Source Code* Ekstraksi Fitur TF-IDF
- Lampiran 14. Operator yang digunakan
- Lampiran 15. Proses dalam PSO + RF
- Lampiran 16. Gambar Hasil *Confusion Matrix* Akurasi *Random Forest* 70:30
- Lampiran 17. Gambar Hasil *Confusion Matrix* Akurasi PSO + *Random Forest* 70:30
- Lampiran 18. Gambar Hasil *Confusion Matrix* Akurasi *Random Forest* 80:20
- Lampiran 19. Gambar Hasil *Confusion Matrix* Akurasi PSO + *Random Forest* 80:20
- Lampiran 20. Gambar Hasil *Confusion Matrix* Akurasi *Random Forest* 90:10
- Lampiran 21. Gambar Hasil *Confusion Matrix* Akurasi PSO + *Random Forest* 90:10