

**OPTIMASI KINERJA ALGORITMA NAIVE BAYES,
DECISION TREE, DAN K-NEAREST NEIGHBORS (KNN)
MENGUNAKAN PARTICLE SWARM OPTIMIZATION
(PSO) DALAM KLASIFIKASI KEPUASAN PELAYANAN
KESEHATAN**

SKRIPSI

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



disusun oleh

RIO EFRANDI

20.11.3660

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

**OPTIMASI KINERJA ALGORITMA NAIVE BAYES, DECISION TREE,
DAN K-NEAREST NEIGHBORS (KNN) MENGGUNAKAN PARTICLE
SWARM OPTIMIZATION (PSO) DALAM KLASIFIKASI KEPUASAN
PELAYANAN KESEHATAN**

HALAMAN JUDUL

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana

Program Studi informatika



disusun oleh

RIO EFRANDI

20.11.3660

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2025**

HALAMAN PERSETUJUAN

SKRIPSI

OPTIMASI KINERJA ALGORITMA NAIVE BAYES, DECISION TREE, DAN K- NEAREST NEIGHBORS (KNN) MENGGUNAKAN PARTICLE SWARM OPTIMIZATION (PSO) DALAM KLASIFIKASI KEPUASAN PELAYANAN KESEHATAN

yang disusun dan diajukan oleh

Rio Efrandi

20.11.3660

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 24 Februari 2025

Dosen Pembimbing,



Windha Mega Pradnya Duhita, S.Kom., M.Kom
NIK. 190302185

HALAMAN PENGESAHAN

SKRIPSI

OPTIMASI KINERJA ALGORITMA NAIVE BAYES, DECISION TREE, DAN K-NEAREST NEIGHBORS (KNN) MENGGUNAKAN PARTICLE SWARM OPTIMIZATION (PSO) DALAM KLASIFIKASI KEPUASAN PELAYANAN KESEHATAN

yang disusun dan diajukan oleh

Rio Efrandi

20.11.3660

Telah dipertahankan di depan Dewan Penguji
pada tanggal 24 Februari 2025

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Theopilus Bayu Sasongko, S.Kom., M.Eng.
NIK. 190302375



Rumini, S.Kom., M.Kom.
NIK. 190302246



Windha Mega Pradnya Dhuhiita, S.Kom., M.Kom.
NIK. 190302185



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 24 Februari 2025

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusriini, M.Kom.
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama Mahasiswa : Rio Efrandi
NIM : 20.11.3660

Menyatakan bahwa Skripsi dengan judul berikut:

OPTIMASI KINERJA ALGORITMA NAIVE BAYES, DECISION TREE, DAN K-NEAREST NEIGHBORS (KNN) MENGGUNAKAN PARTICLE SWARM OPTIMIZATION (PSO) DALAM KLASIFIKASI KEPUASAN PELAYANAN KESEHATAN

Dosen Pembimbing : Windha Mega Pradnya Dhuhiita, S.Kom., 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, 24 Februari 2025

Yang Menyatakan,



Rio Efrandi

HALAMAN PERSEMBAHAN

Puji syukur saya panjatkan ke hadirat Allah SWT atas segala rahmat dan karunia-Nya, sehingga saya dapat menyelesaikan skripsi ini dengan baik. Skripsi ini disusun sebagai syarat menyelesaikan program sarjana prodi Informatika Universitas Amikom Yogyakarta.

Serta saya persembahkan juga kepada:

1. Ayah dan Ibu tercinta, yang dengan kasih sayang, doa, dan dukungan yang tak pernah putus, menjadi sumber kekuatan terbesar dalam hidup saya. Terima kasih atas segala pengorbanan, semangat, dan cinta yang tiada tara hingga saya mampu menyelesaikan skripsi ini.
2. Saya persembahkan kepada Rio Efrandi yaitu diri saya sendiri karena telah berjuang sampai akhir, tidak sedikit halangan rintangan dalam hidup yang dilalui tanpa menyerah
3. Ibu Windha Mega Pradnya D, M.Kom selaku dosen pembimbing di program studi yang telah membagikan ilmu, membimbing, dan memberikan arahan selama masa skripsi.
4. Teman-teman seperjuangan yang telah menjadi tempat berbagi cerita, saling menyemangati, dan tumbuh bersama.
5. Dan semua pihak yang telah memberikan dukungan, doa, dan semangat baik secara langsung maupun tidak langsung dalam proses ini.

Semoga karya ini menjadi langkah awal untuk terus belajar, berkembang, dan bermanfaat bagi banyak orang.

Yogyakarta, 24 Februari 2025



Rio Efrandi

KATA PENGANTAR

Alhamdulillah, segala puji bagi Allah SWT atas segala rahmat dan kemudahan-Nya, sehingga penulis dapat menyelesaikan skripsi ini yang berjudul: **"OPTIMASI KINERJA ALGORITMA NAIVE BAYES, DECISION TREE, DAN K-NEAREST NEIGHBORS (KNN) MENGGUNAKAN PARTICLE SWARM OPTIMIZATION (PSO) DALAM KLASIFIKASI KEPUASAN PELAYANAN KESEHATAN"**

sebagai salah satu syarat menyelesaikan studi di Program Studi Informatika, Universitas Amikom Yogyakarta. Skripsi ini adalah hasil dari proses belajar panjang, penuh tantangan, dan tentu saja banyak momen begadang. Penulis berharap hasil sederhana ini bisa memberikan sedikit kontribusi dalam bidang data mining dan pengolahan data pelayanan kesehatan.

Penulis menyadari masih banyak kekurangan dalam penyusunan skripsi ini. Oleh karena itu, saran dan kritik yang membangun sangat penulis harapkan sebagai bahan perbaikan ke depannya. Semoga skripsi ini bisa bermanfaat, baik untuk penulis sendiri maupun bagi siapa pun yang membacanya.

Yogyakarta, 24 Februari 2025



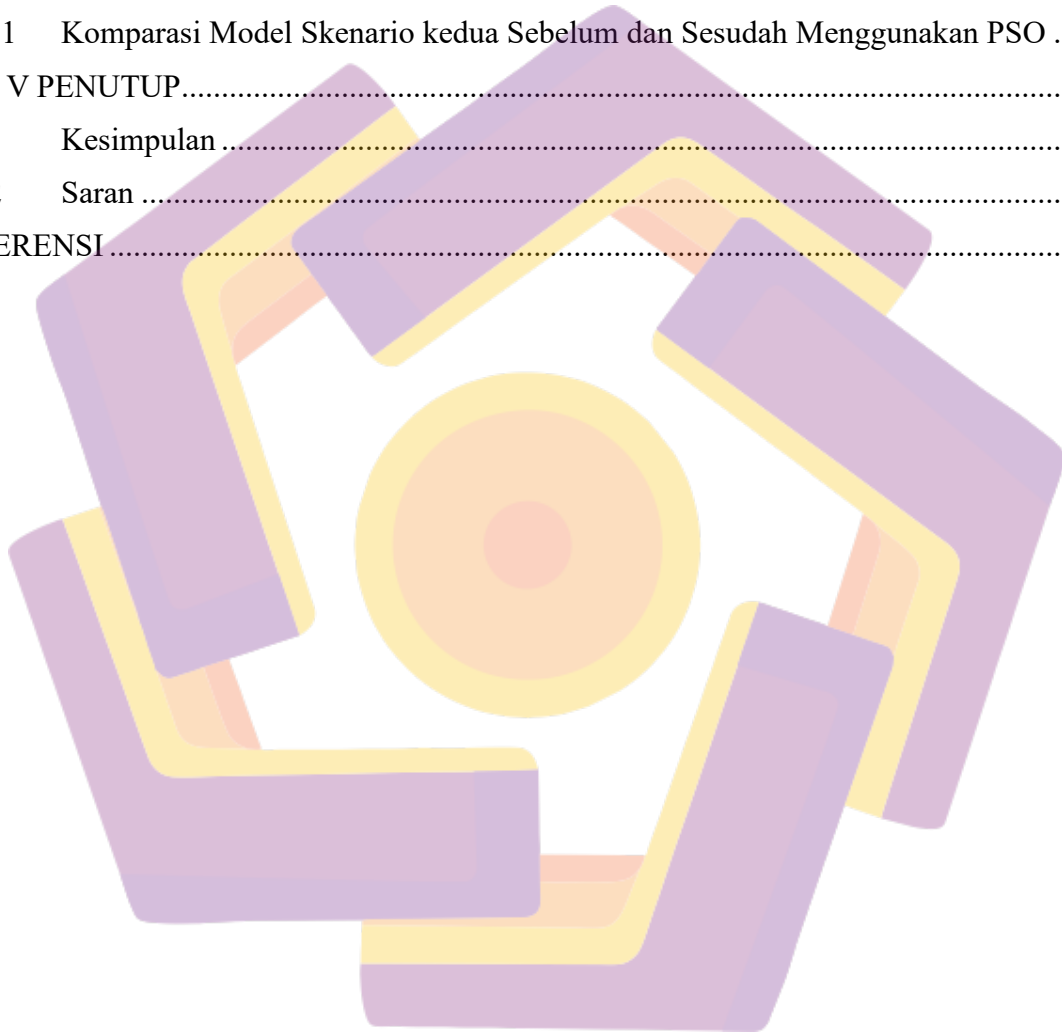
Rio Efrandi

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 LAMBANG DAN SINGKATAN	xii
DAFTAR ISTILAH	xiii
INTISARI	xiv
ABSTRACT.....	xv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah	2
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.6 Sistematika Penulisan	3
BAB II TINJAUAN PUSTAKA.....	5
2.1 Studi Literatur	5
2.2 Dasar Teori.....	10
2.2.1 Pelayanan Kesehatan	10
2.2.2 Web Crawling	11
2.2.3 Sentiment Analisis	11
2.2.4 Komparasi	11
2.2.5 Exploratory Data Analysis	12
2.2.6 Algoritma naive bayes	12
2.2.7 Algoritma decision tree.....	13
2.2.8 Algoritma KNN	14

2.2.9	PSO	14
2.2.10	Confusion Matrix	15
BAB III METODE PENELITIAN		18
3.1	Objek Penelitian.....	18
3.2	Alur Penelitian	18
3.3	Pengumpulan Dataset Pelayanan Kesehatan	18
3.4	Labeling	19
3.5	Preprocessing	19
3.5.1	Cleaning	19
3.5.2	Case Folding	19
3.5.3	Tokenizing dan Stopword Removal	19
3.5.4	Stemming	20
3.6	EDA (Exploratory Data Analysis)	20
3.7	Pembagian Data	20
3.8	Vectorization (Representasi Fitur)	21
3.9	SMOTE	21
3.10	Pelatihan Model	21
3.11	Optimasi Model Menggunakan PSO	22
3.12	Evaluasi.....	22
3.13	Alat dan Bahan.....	22
BAB IV HASIL DAN PEMBAHASAN		23
4.1	Pengumpulan Data	23
4.2	Pelabelan Data Manual	24
4.3	Hasil Preprocessing.....	25
4.3.1	Cleaning	25
4.3.2	Case Folding	26
4.3.3	Tokenizing Dan Stopword Removal	27
4.3.4	Stemming	29
4.4	EDA (Exploratory Data Analysis)	29
4.4.1	Distribusi Kelas.....	30
4.4.2	Grafik Frekuensi Kata.....	30
4.5	Pembagian Data	31
4.6	Vectorization (Representasi Fitur)	32
4.7	SMOTE (Synthetic Minority Over-sampling Technique).....	33
4.8	Hasil Klasifikasi.....	34

4.8.1	Naive bayes	35
4.8.2	Decision Tree	37
4.8.3	K-Nearest Neighbor	38
4.9	Komparasi Model Naïve Bayes, Decision Tree, dan KNN	39
4.10	Hasil Pelatihan Model Sesudah Di Optimasi.....	39
4.10.1	Naive bayes	41
4.10.2	Decision Tree	42
4.10.3	K-Nearest Neighbor	44
4.11	Komparasi Model Skenario kedua Sebelum dan Sesudah Menggunakan PSO	45
BAB V PENUTUP.....		48
5.1	Kesimpulan	48
5.2	Saran	49
REFERENSI		50



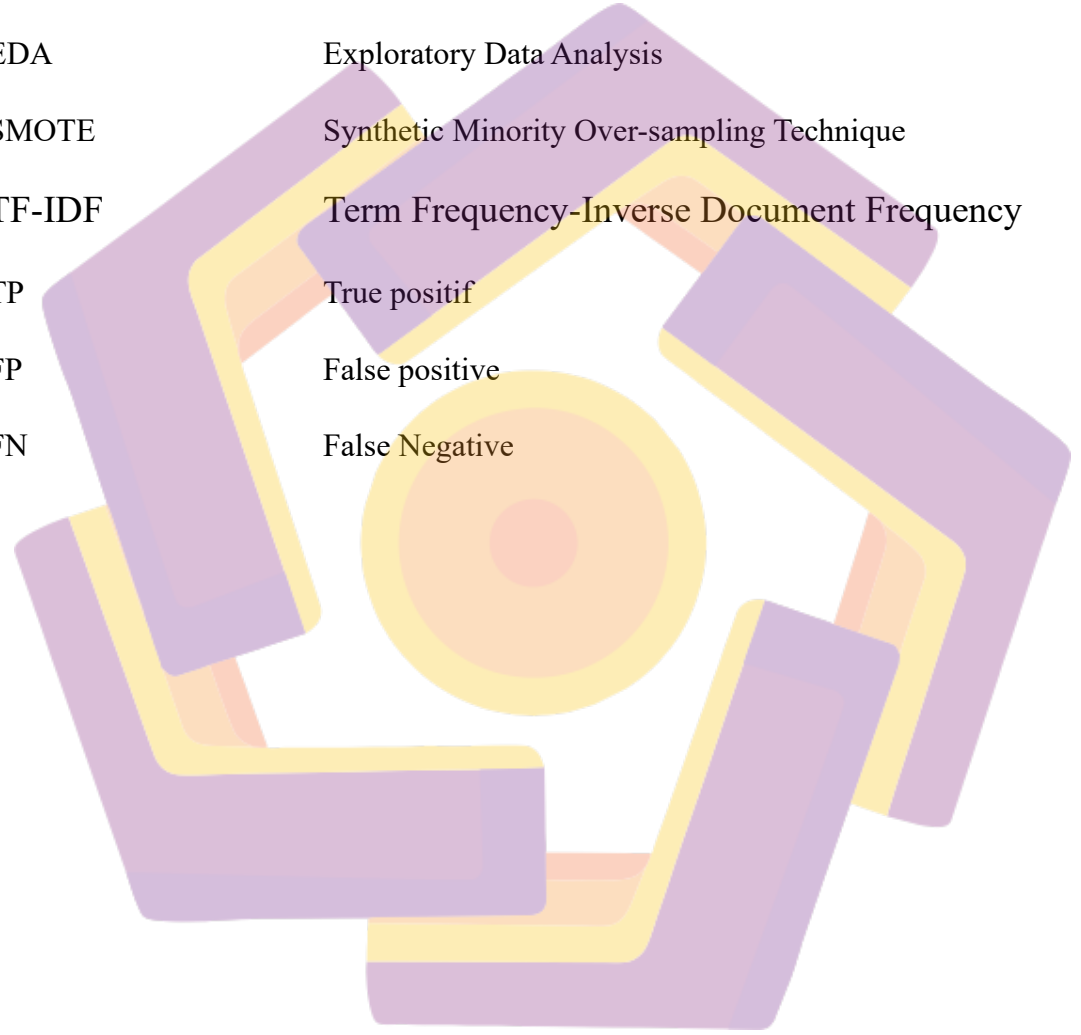
DAFTAR TABEL

Tabel 2.1 Keaslian Penelitian	7
Tabel 2.2 Ilustrasi Confusion Matrix	16
Table 4.1 Pelabelan Manual	24
Table 4.2 Jumlah Data	24
Table 4.3 Hasil Pelabelan Manual	24
Table 4.4 Hasil Cleaning	25
Table 4.5 Hasil Case folding	26
Table 4.6 Hasil Tokenizing Dan Stopword Removal	28
Table 4.7 Hasil Stemming	29
Table 4.8 Hasil Pembagian Data	31
Table 4.9 Distribusi Label Sentimen Data Training Dan Testing	32
Tabel 4.10 Perbandingan Data Sebelum dan Sesudah SMOTE	34
Table 4.11 Hasil Skenario Pertama Tanpa di Optimasi	34
Table 4.12 Hasil Skenario Kedua Setelah di Optimasi	39
Tabel 4.13 hasil perhitungan Euclidean Distance	45
Table 4.14 Hasil Komparasi Sebelum Dan Sesudah PSO	46

DAFTAR GAMBAR

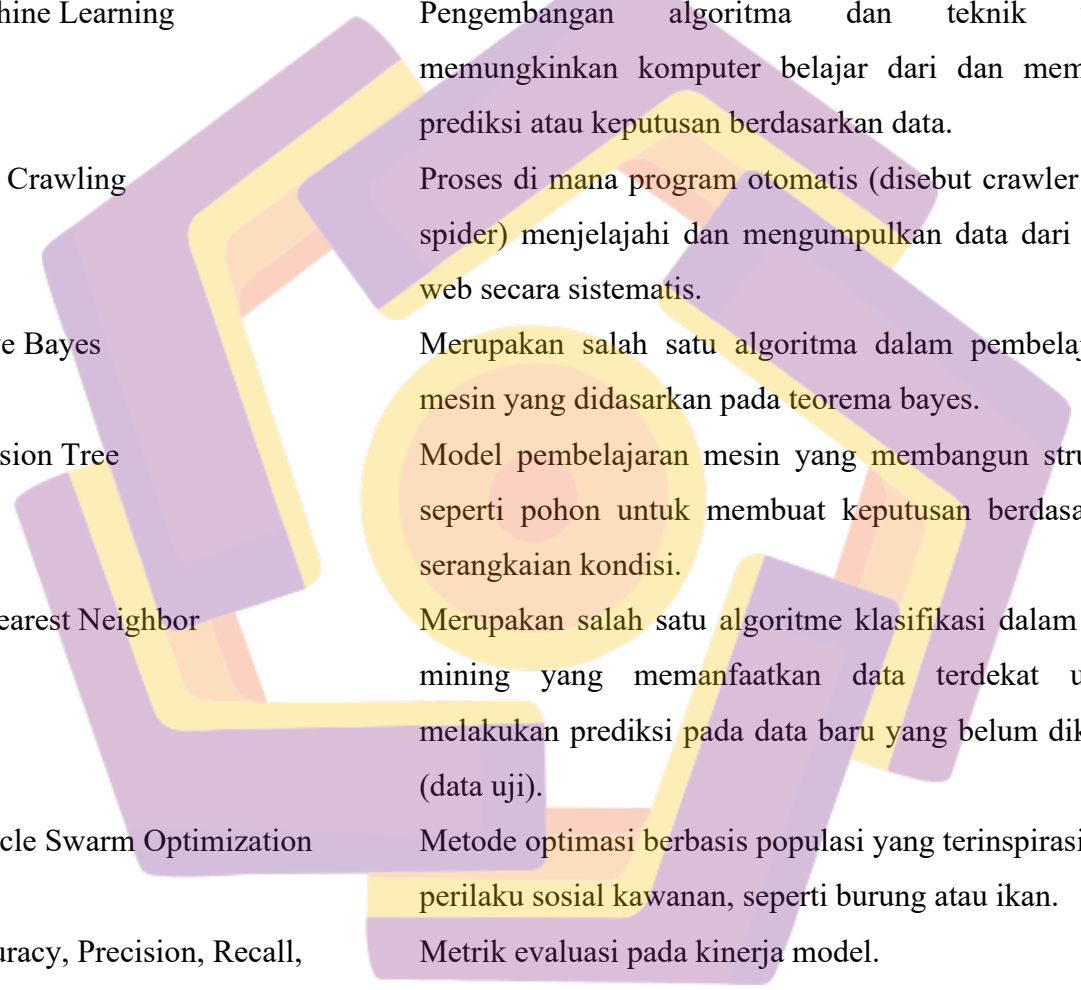
Gambar 3.1 Alur Peneltian	16
Gambar 4.1 Crawling Dataset	21
Gambar 4.2 Distribusi Kelas Positif dan Negatif	28
Gambar 4.3 20 kata yang paling sering muncul	29
Gambar 4.4 hasil vectorization	30
Gambar 4.5 Kelas Diagram Sesudah SMOTE	31
Gambar 4.6 Kelas Diagram Sebelum SMOTE	31
Gambar 4.7 Evaluasi Confusion Matrix Naive Bayes Sebelum di Optimasi	33
Gambar 4.8 Evaluasi Confusion Matrix Decision Tree Sebelum di Optimasi	34
Gambar 4.9 Evaluasi Confusion Matrix KNN Sebelum di Optimasi	35
Gambar 4.10 Hasil Komparasi Akurasi Persentase ke tiga Model	36
Gambar 4.11 Hasil Confusion Matrix Naive Bayes setelah di optimasi	38
Gambar 4.12 Hasil Confusion Matrix Decision Tree setelah di optimasi	39
Gambar 4.13 Hasil Confusion Matrix KNN setelah di optimasi	40
Gambar 4.14 Hasil Komparasi Akurasi Persentase ke tiga Model sebelum dan sesudah di optimasi.	41

DAFTAR LAMBANG DAN SINGKATAN



KNN	k-Nearest Neighbor
PSO	Particle Swarm Optimization
BPJS	Badan Penyelenggara Jaminan Sosial
NLP	Natural Language Processing
EDA	Exploratory Data Analysis
SMOTE	Synthetic Minority Over-sampling Technique
TF-IDF	Term Frequency-Inverse Document Frequency
TP	True positif
FP	False positive
FN	False Negative

DAFTAR ISTILAH

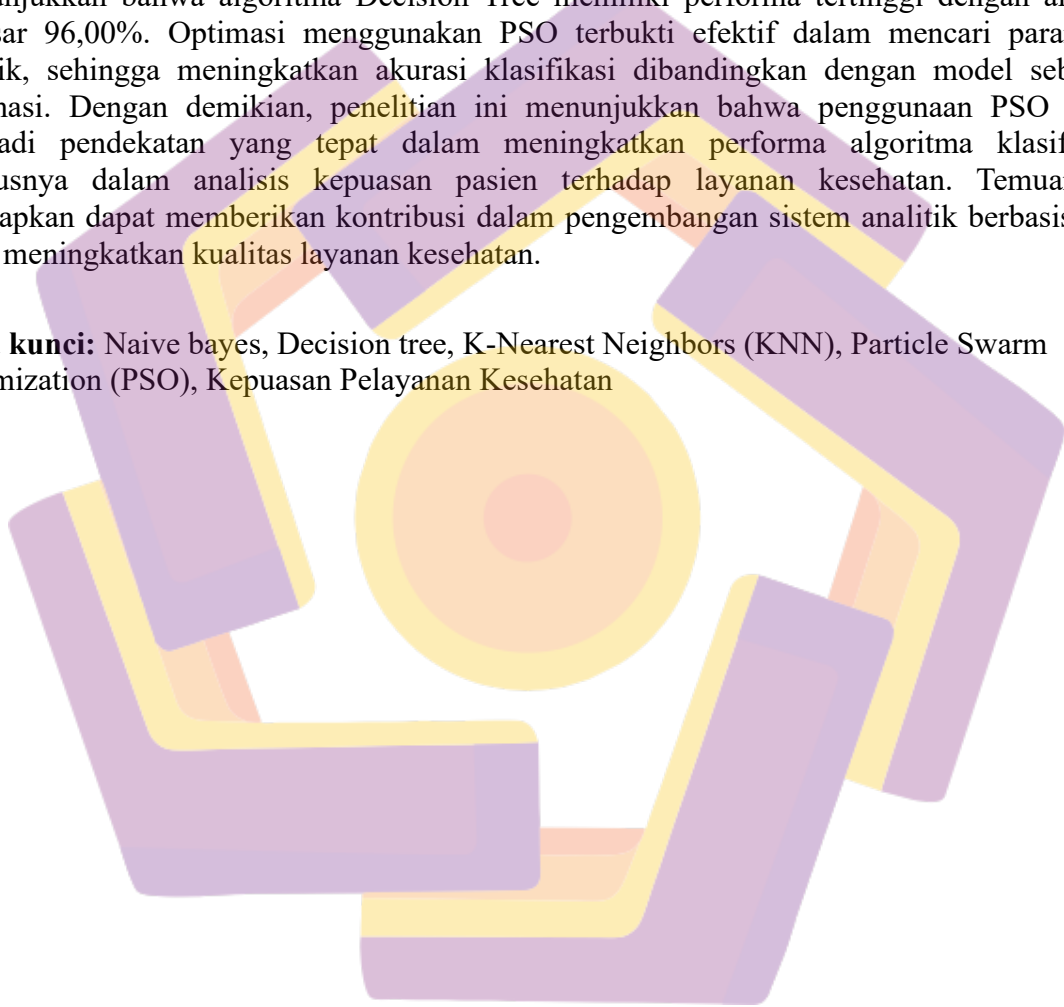


X (Twitter)	Platform media sosial dan layanan jejaring sosial yang dioperasikan oleh X Corp. Perubahan nama ini terjadi pada Juli 2023 sebagai bagian dari visi besar Elon Musk menciptakan aplikasi super.
Ulasan	Penilaian, pendapat, atau pengalaman pribadi tentang suatu tempat, produk, atau layanan.
Machine Learning	Pengembangan algoritma dan teknik yang memungkinkan komputer belajar dari dan membuat prediksi atau keputusan berdasarkan data.
Web Crawling	Proses di mana program otomatis (disebut crawler atau spider) menjelajahi dan mengumpulkan data dari situs web secara sistematis.
Naive Bayes	Merupakan salah satu algoritma dalam pembelajaran mesin yang didasarkan pada teorema bayes.
Decision Tree	Model pembelajaran mesin yang membangun struktur seperti pohon untuk membuat keputusan berdasarkan serangkaian kondisi.
K-Nearest Neighbor	Merupakan salah satu algoritme klasifikasi dalam data mining yang memanfaatkan data terdekat untuk melakukan prediksi pada data baru yang belum dikenal (data uji).
Particle Swarm Optimization	Metode optimasi berbasis populasi yang terinspirasi dari perilaku sosial kawanan, seperti burung atau ikan.
Accuracy, Precision, Recall, dan F1-Score	Metrik evaluasi pada kinerja model.

INTISARI

Kesehatan masyarakat merupakan aspek penting dalam pembangunan suatu negara, dan kepuasan pelayanan kesehatan menjadi indikator utama dalam menilai efektivitas sistem kesehatan. Penelitian ini bertujuan untuk menganalisis performa algoritma Naïve Bayes, Decision Tree, dan K-Nearest Neighbors (KNN) sebelum dan setelah dioptimalkan menggunakan Particle Swarm Optimization (PSO) dalam klasifikasi kepuasan pelayanan kesehatan. Dataset yang digunakan dikumpulkan dari platform X (Twitter) dengan total 2.630 tweet, terdiri dari 1.587 data positif dan 860 data negatif. Data tersebut melalui tahap preprocessing untuk pembersihan dan persiapan sebelum pelatihan model. Hasil penelitian menunjukkan bahwa algoritma Decision Tree memiliki performa tertinggi dengan akurasi sebesar 96,00%. Optimasi menggunakan PSO terbukti efektif dalam mencari parameter terbaik, sehingga meningkatkan akurasi klasifikasi dibandingkan dengan model sebelum optimasi. Dengan demikian, penelitian ini menunjukkan bahwa penggunaan PSO dapat menjadi pendekatan yang tepat dalam meningkatkan performa algoritma klasifikasi, khususnya dalam analisis kepuasan pasien terhadap layanan kesehatan. Temuan ini diharapkan dapat memberikan kontribusi dalam pengembangan sistem analitik berbasis data guna meningkatkan kualitas layanan kesehatan.

Kata kunci: Naive bayes, Decision tree, K-Nearest Neighbors (KNN), Particle Swarm Optimization (PSO), Kepuasan Pelayanan Kesehatan



ABSTRACT

Public health is an important aspect of a country's development, and health service satisfaction is a key indicator in assessing the effectiveness of the health system. This study aims to analyze the performance of Naïve Bayes, Decision Tree, and K-Nearest Neighbors (KNN) algorithms before and after being optimized using Particle Swarm Optimization (PSO) in the classification of health service satisfaction. The dataset used was collected from platform X (Twitter) with a total of 2,630 tweets, consisting of 1,587 positive data and 860 negative data. The data went through a preprocessing stage for cleaning and preparation before model training. The results showed that the Decision Tree algorithm had the highest performance with an accuracy of 96.00%. Optimization using PSO proved effective in finding the best parameters, thus improving classification accuracy compared to the model before optimization. Thus, this study shows that the use of PSO can be an appropriate approach in improving the performance of classification algorithms, especially in analyzing patient satisfaction with health services. The findings are expected to contribute to the development of data-driven analytics systems to improve the quality of healthcare services.

Keyword: *Naive bayes, Decision tree, K-Nearest Neighbors (KNN), Particle Swarm Optimization (PSO), Health Service Satisfaction*

