



IMPLEMENTASI ADAPTIVE BOOSTING DENGAN SELEKSI FITUR

PSO PADA KLASIFIKASI DATA PENYAKIT STROKE

Skripsi

Untuk Memenuhi Persyaratan

Dalam Menyelesaikan Sarjana Strata-1 Ilmu Komputer

Oleh

MUHAMMAD ALI AKBAR

NIM. 1711016210014

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

APRIL 2024



**IMPLEMENTASI ADAPTIVE BOOSTING DENGAN SELEKSI FITUR
PSO PADA KLASIFIKASI DATA PENYAKIT STROKE**

Skripsi

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

**MUHAMMAD ALI AKBAR
NIM. 1711016210014**

**PROGRAM STUDI S-1 ILMU KOMPUTER
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMBUNG MANGKURAT
BANJARBARU
APRIL 2024**

SKRIPSI

IMPLEMENTASI ADAPTIVE BOOSTING DENGAN SLEKSI FITUR PSO PADA KLASIFIKASI DATA PENYAKIT STROKE.

Oleh :

MUHAMMAD ALI AKBAR

NIM. 1711016210014

Telah dipertahankan di depan Dosen Penguji pada
Susunan Dosen Penguji :

Pembimbing I



Muliadi, S.Kom M.Cs
NIP. 197804222010121002

Dosen Penguji I



Dodon Turianto Nugrahadi, S.Kom, M.Eng
NIP. 198001122009121002

Pembimbing II



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

Dosen Penguji II



Triando Hamonangan Saragih, S.Kom,
M.Kom
NIP. 199308242019031012



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

Irwan 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 diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam Daftar Pustaka.

Banjarbaru, April 2024



Muhammad Ali Akbar
NIM. 1711016210014

ABSTRAK

IMPLEMENTASI ADAPTIVE BOOSTING DENGAN SELEKSI FITUR PSO PADA KLASIFIKASI DATA PENYAKIT STROKE (Oleh: M. Ali Akbar; Pembimbing: Pembimbing 1 dan Pembimbing 2; 2024; 69 halaman)

Stroke, sebuah kondisi neurologis serius, merupakan penyakit yang menyebabkan tingkat kecacatan dan kematian yang tinggi di seluruh dunia. Organisasi Kesehatan Dunia (WHO) mencatat bahwa setiap tahunnya, jutaan orang di seluruh dunia menderita stroke, dan dampaknya tidak hanya terbatas pada kesehatan individu, tetapi juga memiliki implikasi sosial dan ekonomi yang signifikan. Diagnosis dini dan akurat dari stroke memainkan peran kunci dalam menyediakan perawatan yang tepat waktu dan efektif bagi para pasien. Namun, diagnosis stroke seringkali merupakan tantangan, terutama karena gejalanya yang bervariasi dan tumpang tindih dengan kondisi kesehatan lainnya.

Dalam upaya meningkatkan akurasi dan keandalan diagnosis stroke, pendekatan klasifikasi data telah menjadi fokus utama dalam penelitian medis. Dalam konteks ini, penggunaan metode klasifikasi yang canggih, seperti *Adaboost* dan *Particle Swarm Optimization* (PSO), telah menarik perhatian para peneliti. *Adaboost*, atau *Adaptive Boosting*, adalah teknik klasifikasi yang kuat dalam *machine learning*. Metode ini bekerja dengan menggabungkan sejumlah besar "lemah pembelajaran" (*weak learners*), seperti *decision trees*, menjadi model yang kuat. Dengan kemampuannya untuk menangani data yang tidak seimbang dan mengurangi *overfitting*, *Adaboost* telah terbukti efektif dalam banyak aplikasi klasifikasi, termasuk dalam bidang medis.

Di sisi lain, *Particle Swarm Optimization* (PSO) adalah algoritma optimisasi yang terinspirasi oleh perilaku kawanan partikel dalam pencarian makanan. PSO telah digunakan secara luas dalam seleksi fitur, yaitu proses pemilihan subset fitur terbaik dari sekumpulan fitur yang ada. Dengan menggunakan PSO untuk seleksi fitur, kita dapat meningkatkan efisiensi dan kinerja model klasifikasi dengan mengurangi dimensi fitur yang tidak relevan atau redundan, serta meningkatkan interpretabilitas model.

Studi ini bertujuan untuk menguji kinerja model klasifikasi menggunakan metode *Adaboost* dengan dan tanpa seleksi fitur *PSO* pada Dataset Penyakit Stroke. Hasil pengujian menunjukkan bahwa *Adaboost* dengan seleksi fitur *PSO* menghasilkan nilai akurasi yang lebih tinggi daripada *Adaboost* tanpa seleksi fitur. Nilai akurasi tertinggi dicapai oleh model klasifikasi *Adaboost* dengan *PSO*, menunjukkan potensi untuk meningkatkan diagnosis stroke dan memberikan perawatan yang lebih efektif bagi pasien. Dengan demikian, penggunaan kombinasi *Adaboost* dan *PSO* dapat menjadi alat yang berguna dalam bidang medis untuk mengatasi tantangan dalam diagnosis penyakit yang kompleks seperti stroke.

Kata kunci: *Stroke, Klasifikasi, Adaboost, Particle Swarm Optimization*

ABSTRACT

IMPLEMENTATION OF ADAPTIVE BOOSTING WITH PSO FEATURE SELECTION IN STROKE DISEASE DATA CLASSIFICATION (By: M. Ali Akbar; Advisors: Advisor 1 and Advisor 2; 2024; 69 pages)

Stroke, a serious neurological condition, is a disease that causes high rates of disability and death worldwide. The World Health Organization (WHO) notes that millions of people worldwide suffer from stroke each year, with its impact extending beyond individual health to significant social and economic implications. Early and accurate diagnosis of stroke plays a key role in providing timely and effective treatment for patients. However, diagnosing stroke is often challenging, particularly due to its varied symptoms that overlap with other health conditions. In efforts to enhance the accuracy and reliability of stroke diagnosis, data classification approaches have become a primary focus in medical research. In this context, the use of advanced classification methods such as Adaboost and Particle Swarm Optimization (PSO) has garnered researchers' attention. Adaboost, also known as Adaptive Boosting, is a powerful classification technique in machine learning. It works by combining numerous weak learners, such as decision trees, into a robust model. With its ability to handle imbalanced data and reduce overfitting, Adaboost has proven effective in various classification applications, including in the medical field.

On the other hand, Particle Swarm Optimization (PSO) is an optimization algorithm inspired by the behavior of particle swarms in food search. PSO has been widely used in feature selection, the process of selecting the best subset of features from a set of existing features. By using PSO for feature selection, we can enhance the efficiency and performance of classification models by reducing irrelevant or redundant feature dimensions, as well as improving model interpretability.

This study aims to evaluate the performance of classification models using the Adaboost method with and without PSO feature selection on a Stroke Disease Dataset. The results of the testing indicate that Adaboost with PSO feature selection yields higher accuracy compared to Adaboost without feature selection. The highest accuracy is achieved by the Adaboost classification model with PSO, demonstrating the potential to improve stroke diagnosis and provide more effective treatment for patients. Thus, the combined use of Adaboost and PSO can be a valuable tool in the medical field to address challenges in diagnosing complex diseases such as stroke.

Keywords: Stroke, Classification, Adaboost, Particle Swarm Optimization

PRAKATA

Assalamualaikum Warahmatullahi Wabarakatuh.

Puji syukur penulis panjatkan ke hadirat Allah SWT karena atas berkat rahmat dan karunia-Nya penulis dapat menyelesaikan skripsi yang berjudul “IMPLEMENTASI ADAPTIVE BOOSTING DENGAN SELEKSI FITUR PSO PADA KLASIFIKASI DATA PENYAKIT STROKE” untuk memenuhi syarat dalam menyelesaikan pendidikan program S1 Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lambung Mangkurat. Tak lupa pula penulis panjatkan sholawat dan salam ke hadirat Rasulullah Muhammad SAW beserta para sahabat, keluarga, dan pengikut beliau hingga *yaumul qiamah*.

Pada lembar ini penulis ingin menyampaikan ucapan terimakasih kepada pihak-pihak yang sangat mendukung penulis dalam pembuatan dan penyusunan skripsi ini, adapun yang dimaksud adalah sebagai berikut :

Akhir kata penulis menyadari sepenuhnya bahwa penulisan ini jauh dari sempurna, namun penulis mengharapkan bantuan serupa 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, April 2024



Penulis

DAFTAR ISI

LEMBAR PENGESAHAN SKRIPSI.....	iii
LEMBAR PERNYATAAN.....	iv
ABSTRAK.....	v
<i>ABSTRACT</i>	vi
PRAKATA.....	vii
DAFTAR ISI.....	viii
DAFTAR TABEL.....	x
DAFTAR GAMBAR.....	xi
DAFTAR LAMPIRAN.....	xii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan Penelitian	3
1.4 Manfaat Penelitian	3
1.5 Batasan Masalah	3
BAB II TINJAUAN PUSTAKA	5
2.1 Kajian Terdahulu	5
2.2 Penyakit Stroke	7
2.3 Faktor Penyebab Stroke	9
2.4 Gejala Penyakit Stroke	10
2.5 Data	11
2.6 <i>Data Mining</i>	12
2.7 <i>Arsitektur Data Mining</i>	16
2.8 <i>Data Preprocessing</i>	17
2.9 Klasifikasi	18
2.10 <i>Particle Swarm Optimization (PSO)</i>	19
2.11 <i>Adaptive Boosting (Adaboost)</i>	20
2.12 <i>Confusion Matrix</i>	21
2.13 Median	22

2.14	Teknik <i>Ensemble</i>	24
2.15	<i>K-Fold Cross Validation</i>	24
2.16	Phyton	25
BAB III METODE PENELITIAN		26
3.1	Alat Penelitian	26
3.2	Bahan Penelitian	26
3.3	Variabel Penelitian	27
3.4	Prosedur Penelitian	28
3.4.1	Pengumpulan Data	32
3.4.2	Pengolahan Data	33
3.4.3	Pelatihan Model.....	34
3.4.4	Pengujian Model Klasifikasi	37
3.4.5	Evaluasi dan Analisis Hasil.....	37
BAB IV HASIL DAN PEMBAHASAN.....		38
4.1	Hasil	38
4.1.1	Pengumpulan Data	38
4.1.2	Pengolahan Data	40
4.1.3	Pembuatan Model Klasifikasi.....	43
4.1.4	Klasifikasi dan Evaluasi.....	46
4.2	Analisis Hasil	48
4.3	Pembahasan	49
BAB V PENUTUP.....		51
5.1	Kesimpulan	51
5.2	Saran	51
DAFTAR PUSTAKA		52

DAFTAR TABEL

Tabel 1. Penelitian Terdahulu	6
Tabel 2. Rancangan Penelitian	7
Tabel 3. Confusion Matrix.....	22
Tabel 4. Data Contoh Penanganan Missing Value dengan Nilai Median	23
Tabel 5. Deskripsi Atribut Penyakit Stroke.....	32
Tabel 6. Skenario Pelatihan Model Klasifikasi	34
Tabel 7. Dataset Penyakit Stroke.....	39
Tabel 8. Data Cleaning.....	41
Tabel 9. Hasil Pengujian Nilai Akurasi Model Klasifikasi Adaboost Menggunakan 10-fold cross validation.....	46
Tabel 10. Hasil Pengujian Nilai Akurasi Model Klasifikasi Adaboost + PSO Menggunakan 10-fold cross validation.....	47

DAFTAR GAMBAR

Gambar 1. Langkah-langkah dalam proses data mining	13
Gambar 2. Arsitektur Data Mining.....	17
Gambar 3. K-Fold Cross Validation.....	25
Gambar 4. Diagram Alir Proses Penelitian (Skenario 1).....	30
Gambar 5. Diagram Alir Proses Penelitian (Skenario 2).....	31
Gambar 6. Flowchart Skenario Klasifikasi.....	35
Gambar 7. Persentase Data Pasien Stroke VS Pasien Non Stroke	40

DAFTAR LAMPIRAN

Lampiran 1. Source Code Model Klasifikasi Adaboost	57
Lampiran 2. Source Code Model Klasifikasi Adaboost + PSO	63