

**OPTIMASI ALGORITMA *RANDOM FOREST* MENGGUNAKAN
ALGORITMA GENETIKA UNTUK KLASIFIKASI DATA PENDERITA
PENYAKIT STROKE**

Diajukan Sebagai Syarat Untuk Menyelesaikan
Pendidikan Program Strata-1 Pada
Jurusan Teknik Informatika



Oleh:

Citra Meida Antika
NIM: 09021182025007

Jurusan Teknik Informatika
FAKULTAS ILMU KOMPUTER UNIVERSITAS SRIWIJAYA
2024

LEMBAR PENGESAHAN SKRIPSI

**OPTIMASI ALGORITMA *RANDOM FOREST* MENGGUNAKAN
ALGORITMA GENETIKA UNTUK KLASIFIKASI DATA PENDERITA
PENYAKIT STROKE**

Oleh:

CITRA MEIDA ANTIKA

NIM: 09021182025007

Palembang, 13 Juni 2024

Mengetahui,
Ketua Jurusan Teknik Informatika



Dr. M. Fachrurrozi, S.Si., M.T.
NIP 198005222008121002

Pembimbing



Dian Palupi Rini M.Kom., Ph.D.
NIP 197802232006042002

TANDA LULUS UJIAN KOMPREHENSIF SKRIPSI

Pada hari tanggal 5 Juni 2024 telah dilaksanakan ujian komprehensif skripsi oleh
Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya

Nama : Citra Meida Antika
NIM : 09021182025007
Judul : Optimasi Algoritma *Random Forest* Menggunakan Algoritma
Genetika Untuk Klasifikasi Data Penderita Penyakit Stroke

dan dinyatakan LULUS.

1. Ketua Penguji

Kanda Januar Misrawan, M.T.
NIP 199001092019031012



2. Penguji I

Novi Yusliani, M.T.
NIP 198211082012122001



3. Pembimbing I

Dian Palupi Rini, M.Kom., Ph.D.
NIP 197802232006042002



Mengetahui,

Ketua Jurusan Teknik Informatika



Dr. M. Fachrudrozi, S.Si., M.T.
NIP 198005222008121002

HALAMAN PERNYATAAN

Yang bertanda tangan dibawah ini:

Nama : Citra Meida Antika
NIM : 09021182025007
Program Studi : Teknik Informatika
Judul : Optimasi Algoritma *Random Forest* Menggunakan Algoritma Genetika Untuk Klasifikasi Data Penderita Penyakit Stroke

Hasil pengecekan iThenticate/Turnitin: 17%

Menyatakan bahwa Laporan Projek saya merupakan hasil karya sendiri dan bukan hasil penjiplakan/plagiat. Apabila ditemukan unsur penjiplakan/plagiat dalam laporan ini, maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya sesuai dengan ketentuan yang berlaku.

Demikian pernyataan ini saya buat dengan sebenar-benarnya dan tidak ada paksaan oleh siapapun.



Palembang, 13 Juni 2024



Citra Meida Antika
NIM 09021182025007

MOTTO DAN PERSEMBAHAN

”Indeed, with hardship (will be) ease”

Q.S Al-Insyirah: 6

Kupersembahkan karya tulis ini kepada:

- Allah SWT
- Baba dan Ibu
- Adik-adik
- Keluarga Besar
- Dosen Pembimbing
- Almamater

**OPTIMIZATION OF RANDOM FOREST ALGORITHM USING
GENETIC ALGORITHM FOR CLASSIFICATION OF STROKE DISEASE
PATIENT DATA**

By:

Citra Meida Antika (09021182025007)

Department of Informatics, Faculty of Computer Science, Sriwijaya University

Email: citrameiikaa@gmail.com

ABSTRACT

Stroke is the second leading cause of death globally. A lack of knowledge regarding stroke risk factors results in high incidence and mortality rates. Using a classification approach can effectively address this issue. Classification involves grouping categories or classes based on patterns or characteristics found in the data. By classifying stroke patient data, we can identify patterns of stroke risk factors. The Random Forest algorithm was chosen for its high accuracy, but it has a drawback in that it requires a long processing time due to the large number of decision trees that need to be combined to determine the class. To overcome this weakness, a Genetic Algorithm is used for feature selection to identify the most relevant attributes, making the decision tree construction process more efficient and accurate. The study results show that the Random Forest method without optimization achieves an accuracy of 96.4%, while the method optimized with the Genetic Algorithm achieves an accuracy of 97.88%. With resampling, the accuracy of Random Forest reaches 80%, and with Genetic Algorithm optimization, it increases to 84%.

Keywords: *Classification, Genetic Algorithm, Random Forest, Stroke*

Palembang, 13th June 2024



Approved,

Head of Informatics Department



Dr. M. Fachrurrozi, S.Si., M.T.
NIP 198005222008121002

Supervisor



Dian Palupi Kini, M.Kom., Ph.D.
NIP 197802232006042002

OPTIMASI ALGORITMA *RANDOM FOREST* MENGGUNAKAN ALGORITMA GENETIKA UNTUK KLASIFIKASI DATA PENDERITA PENYAKIT STROKE

Oleh:

Citra Meida Antika (09021182025007)

Jurusan Teknik Informatika, Fakultas Ilmu Komputer, Universitas Sriwijaya
Email: citrameiikaa@gmail.com

ABSTRAK

Stroke merupakan salah satu penyebab kematian terbanyak kedua di dunia. Minimnya pengetahuan mengenai faktor risiko stroke mengakibatkan tingginya angka kasus dan kematian. Pendekatan dengan metode klasifikasi merupakan solusi efektif untuk mengatasi masalah ini. Klasifikasi adalah proses pengelompokan kategori atau kelas berdasarkan kesamaan pola atau sifat pada setiap data. Dengan mengklasifikasi data penderita stroke, pola-pola faktor risiko stroke dapat diidentifikasi. Algoritma *Random Forest* dipilih karena kemampuannya menghasilkan akurasi tinggi, namun memiliki kelemahan dalam waktu pemrosesan data yang lama akibat banyaknya pohon keputusan yang harus digabungkan untuk menentukan kelas. Untuk mengatasi kelemahan ini, Algoritma Genetika digunakan untuk seleksi fitur guna mengidentifikasi atribut yang paling relevan, sehingga proses pembuatan pohon keputusan menjadi lebih efisien dan akurat. Hasil penelitian menunjukkan bahwa metode *Random Forest* tanpa optimasi menghasilkan akurasi sebesar 96,4%, sementara metode yang dioptimasi dengan Algoritma Genetika menghasilkan akurasi 97,88%. Dengan *resampling*, akurasi *Random Forest* mencapai 80%, dan dengan optimasi Algoritma Genetika meningkat menjadi 84%.

Kata kunci: Algoritma Genetika, Klasifikasi, *Random Forest*, Stroke

Palembang, 13 Juni 2024

Mengetahui,
Ketua Jurusan Teknik Informatika



Dr. M. Fachrudrozi, S.Si., M.T.
NIP 198005222008121002

Dosen Pembimbing



Dian Palupi Rini, M.Kom., Ph.D.
NIP 197802232006042002

KATA PENGANTAR

Puji syukur penulis panjatkan kehadiran Allah SWT, yang dengan berkat dan rahmat-Nya, penulis berhasil menyelesaikan Tugas Akhir ini. Tugas akhir ini disusun untuk memenuhi salah satu syarat dalam menyelesaikan pendidikan program Strata-1 pada Fakultas Ilmu Komputer, Program Studi Teknik Informatika di Universitas Sriwijaya.

Proses penulisan laporan tugas akhir ini tidak terlepas dari bimbingan, bantuan, dukungan, serta petunjuk dari berbagai pihak. Oleh karena itu, penulis ingin menyampaikan ucapan terima kasih yang tulus kepada:

1. Allah SWT, yang telah memberikan nikmat dan petunjuk-Nya.
2. Seluruh keluarga besar yang telah memberikan dukungan dan doa yang tiada henti.
3. Prof. Dr. Erwin, S.Si., M.Si., selaku Dekan Fakultas Ilmu Komputer.
4. Dr. M. Fachrurrozi, S.Si., M.T., selaku Ketua Jurusan Teknik Informatika.
5. Dian Palupi Rini, M.Kom., Ph.D., selaku Dosen Pembimbing yang telah memberikan bimbingan, arahan, dan motivasi untuk menyelesaikan skripsi.
6. Samsuryadi, M.Kom, Ph.D., selaku Dosen Pembimbing Akademik yang telah memberikan bimbingan dan arahan selama perkuliahan.
7. Novi Yusliani, M.T., selaku Dosen Penguji yang telah memberikan saran dan masukan yang sangat berharga.
8. Seluruh Dosen Jurusan Teknik Informatika dan Dosen Fakultas Ilmu Komputer, atas ilmu yang telah diberikan selama masa perkuliahan.

9. Seluruh Staf Tata Usaha dan Pegawai Fakultas Ilmu Komputer.
10. Betha, Dini, Kurnia, Gading, Tiara, dan Melsra, selaku teman seperjuangan penulis yang telah banyak membantu selama menjalani perkuliahan. Serta pihak-pihak yang tidak bisa penulis sebutkan satu persatu, terima kasih telah melengkapi cerita perjalanan kuliah.

Penulis berharap Tugas Akhir ini dapat menjadi referensi penelitian dan pembelajaran di masa depan. Penulis juga menyadari adanya kekurangan dalam penyusunan Tugas Akhir ini dan mengharapkan kritik dan saran yang membangun untuk meningkatkan kualitas penelitian di masa mendatang. Terakhir, semoga Tugas Akhir ini memberikan manfaat dan kontribusi positif bagi kita semua.

Palembang, 13 Juni 2024

Penulis

DAFTAR ISI

LEMBAR PENGESAHAN SKRIPSI	ii
TANDA LULUS UJIAN KOMPREHENSIF SKRIPSI.....	iii
HALAMAN PERNYATAAN	iv
MOTTO DAN PERSEMBAHAN.....	v
ABSTRACT.....	vi
ABSTRAK.....	vii
KATA PENGANTAR	viii
DAFTAR ISI.....	x
DAFTAR TABEL	xiii
DAFTAR GAMBAR	xiv
BAB I PENDAHULUAN.....	I-1
1.1 Pendahuluan	I-1
1.2 Latar Belakang	I-1
1.3 Rumusan Masalah	I-4
1.4 Tujuan Penelitian	I-4
1.5 Manfaat Penelitian	I-4
1.6 Batasan Masalah	I-5
1.7 Sistematika Penulisan	I-5
1.8 Kesimpulan	I-6
BAB II KAJIAN LITERATUR	II-1
2.1 Pendahuluan	II-1
2.2 Landasan Teori	II-1
2.2.1 Stroke	II-1
2.2.2 Klasifikasi.....	II-2
2.2.3 Decision Tree.....	II-5
2.2.4 Random Forest	II-7
2.2.5 Algoritma Genetika (<i>Genetic Algorithm</i>)	II-9
2.2.6 Pengukuran Akurasi Hasil Klasifikasi.....	II-13
2.2.7 Rational Unified Process	II-13
2.3 Referensi Penelitian Lain	II-14
2.4 Kesimpulan	II-15

BAB III METODOLOGI PENELITIAN	III-1
3.1 Pendahuluan	III-1
3.2 Pengumpulan Data	III-1
3.2.1 Jenis dan Sumber Data	III-1
3.2.2 Metode Pengumpulan Data	III-2
3.3 Tahapan Penelitian	III-2
3.3.1 Kerangka Kerja Penelitian	III-4
3.3.2 Kriteria Pengujian	III-7
3.3.3 Format Data Pengujian	III-8
3.3.4 Alat yang Digunakan Pada Penelitian	III-9
3.3.5 Pengujian Penelitian	III-10
3.3.6 Analisis Hasil Pengujian dan Kesimpulan Penelitian	III-10
3.4 Metode Pengembangan Perangkat Lunak	III-10
3.4.1 Insepsi	III-11
3.4.2 Elaborasi	III-11
3.4.3 Konstruksi	III-11
3.4.4 Transisi	III-11
3.5 Manajemen Proyek Penelitian	III-12
BAB IV PENGEMBANGAN PERANGKAT LUNAK	IV-1
4.1 Pendahuluan	IV-1
4.2 Rational Unified Process (RUP)	IV-1
4.2.1 Fase Insepsi	IV-1
4.2.1.1 Pemodelan Bisnis	IV-1
4.2.1.2 Kebutuhan Sistem	IV-2
4.2.1.3 Analisis dan Perancangan	IV-3
4.2.2 Fase Elaborasi	IV-10
4.2.3 Fase Konstruksi	IV-16
4.2.4 Fase Transisi	IV-20
4.3 Kesimpulan	IV-27
BAB V HASIL DAN ANALISIS PENELITIAN	V-1
5.1 Pendahuluan	V-1
5.2 Data Hasil Percobaan Penelitian	V-1
5.2.1 Konfigurasi Percobaan	V-1
5.2.2 Hasil Pengujian <i>Random Forest</i>	V-2

5.2.3 Hasil Konfigurasi <i>RF+GA</i> Tanpa <i>Resampling</i>	V-3
5.2.4 Hasil Konfigurasi <i>RF+GA</i> Dengan <i>Resampling</i>	V-7
5.3 Analisis Hasil Penelitian	V-8
5.3.1 Hasil Pengujian <i>Random Forest</i>	V-8
5.3.2 Analisis Hasil Pengujian <i>RF+GA</i>	V-10
5.3.4 Hasil Perbandingan <i>RF+GA</i> Dengan <i>Resampling</i>	V-14
5.4 Kesimpulan	V-16
BAB VI	VI-1
KESIMPULAN DAN SARAN	VI-1
6.1 Pendahuluan	VI-1
6.2 Kesimpulan	VI-1
6.3 Saran	VI-2
DAFTAR PUSTAKA	xv
LAMPIRAN	118

DAFTAR TABEL

Tabel III-1. Rancangan Tabel Konfigurasi Pengujian <i>Random Forest</i>	III-5
Tabel III-2. Rancangan Tabel Konfigurasi Pengujian <i>RF+GA</i>	III-5
Tabel III-2. Rancangan Tabel Konfigurasi Pengujian <i>RF+GA resampling</i>	III-5
Tabel III-4. <i>Work Breakdown Structure</i>	III-9
Tabel IV-1. Kebutuhan Fungsional.....	IV-3
Tabel IV-2. Kebutuhan Non Fungsional.....	IV-3
Tabel IV-3. Data Sampel yang akan digunakan.....	IV-4
Tabel IV-4. Definisi Pengguna <i>Use Case</i>	IV-6
Tabel IV-5. Definisi <i>Use Case Diagram</i>	IV-7
Tabel IV-6. Skenario <i>Use Case</i> Melakukan <i>Load Data</i>	IV-7
Tabel IV-7. Skenario <i>Use Case</i> Melakukan Klasifikasi <i>Random Forest</i>	IV-8
Tabel IV-8. Skenario <i>Use Case</i> Melakukan Klasifikasi <i>RF+GA</i>	IV-9
Tabel IV-9. Implementasi Kelas.....	IV-20
Tabel IV-10. Rencana Pengujian <i>Use Case Load Data</i>	IV-24
Tabel IV-11. Rencana Pengujian <i>Use Case</i> Klasifikasi <i>Random Forest</i>	IV-25
Tabel IV-12. Rencana Pengujian <i>Use Case</i> Klasifikasi <i>RF+GA</i>	IV-26
Tabel IV-13. Skenario Pengujian <i>Load Data</i>	IV-27
Tabel IV-14. Skenario Pengujian Klasifikasi <i>Random Forest</i>	IV-27
Tabel IV-15. Skenario Pengujian Klasifikasi <i>RF+GA</i>	IV-29
Tabel V-1. Pengujian Variasi Jumlah Pohon.....	V-2
Tabel V-2. Pengujian Populasi Pada <i>RF+GA</i>	V-3
Tabel V-3. Pengujian Generasi Pada <i>RF+GA</i>	V-4
Tabel V-4. Pengujian <i>Crossover Rate</i> pada <i>RF+GA</i>	V-6
Tabel V-5. Pengujian <i>Mutation Rate</i> Pada <i>RF+GA</i>	V-7
Tabel V-6. Hasil Pengujian <i>RF+GA Resampling</i>	V-8

DAFTAR GAMBAR

Gambar II-1. Struktur <i>Decision Tree</i>	II-3
Gambar II-2. Arsitektur <i>Random Forest</i>	II-6
Gambar II-3. Tahapan Algoritma Genetika.....	II-10
Gambar III-1. Tahapan Penelitian.....	III-3
Gambar III-2. Diagram Kerangka Kerja Penelitian.....	III-3
Gambar III-3. <i>Flowchart</i> Pra-pemrosesan Data.....	III-4
Gambar III-4. <i>Flowchart</i> Klasifikasi.....	III-3
Gambar IV-1. <i>Use Case Diagram</i>	IV-6
Gambar IV-2. Rancangan Antarmuka Klasifikasi <i>Random Forest</i>	IV-12
Gambar IV-3. Rancangan Antarmuka Klasifikasi <i>RF+GA</i>	IV-12
Gambar IV-4. <i>Activity Diagram</i> Load Data.....	IV-14
Gambar IV-5. <i>Activity Diagram</i> Klasifikasi <i>Random Forest</i>	IV-15
Gambar IV-6. <i>Activity Diagram</i> Klasifikasi <i>RF+GA</i>	IV-15
Gambar IV-7. <i>Sequence Diagram</i> Load Data.....	IV-16
Gambar IV-8. <i>Sequence Diagram</i> Klasifikasi <i>Random Forest</i>	IV-17
Gambar IV-9. <i>Sequence Diagram</i> Klasifikasi <i>RF+GA</i>	IV-18
Gambar IV-10. Diagram Kelas.....	IV-20
Gambar IV-11. Antarmuka Klasifikasi <i>Random Forest</i>	IV-22
Gambar IV-12. Antarmuka Klasifikasi <i>RF+GA</i>	IV-2
Gambar V-1. Grafik Hasil Visualisasi <i>Random Forest</i>	V-10
Gambar V-2. Grafik Hasil Visualisasi Berdasarkan Nilai Populasi.....	V-11
Gambar V-3. Grafik Hasil Visualisasi Berdasarkan Nilai Generasi.....	V-12
Gambar V-4. Grafik Hasil Visualisasi Berdasarkan Nilai <i>Crossover Rate</i> ...V-12	
Gambar V-5. Grafik Hasil Visualisasi Berdasarkan Nilai <i>Mutation Rate</i> ...V-13	
Gambar V-6. Grafik Visualisasi Perbandingan <i>RF</i> dan <i>RF+GA</i>	V-14
Gambar V-7. Grafik Visualisasi Perbandingan <i>RF Resampling</i>	V-16
Gambar V-8. Grafik Visualisasi Perbandingan <i>RF+GA Resampling</i>	V-16

BAB I

PENDAHULUAN

1.1 Pendahuluan

Bab ini memberikan penjelasan umum mengenai keseluruhan penelitian, yaitu latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, dan batasan masalah yang menjadi dasar penelitian.

1.2 Latar Belakang

Menurut *World Health Organization* (WHO) pada tahun 2020, stroke merupakan penyakit yang menjadi penyebab kematian terbanyak kedua di dunia setelah penyakit jantung dan menjadi penyebab utama kecacatan nomor tiga di dunia¹. Selain itu, berdasarkan Riskesdas Kementerian Kesehatan Republik Indonesia, stroke menempati posisi tertinggi dalam daftar penyebab kematian di Indonesia pada tahun 2014². Tingginya jumlah kasus stroke dan angka kematian yang disebabkan oleh penyakit ini seringkali disebabkan oleh minimnya pengetahuan masyarakat mengenai faktor risiko stroke.

Oleh karena itu, diperlukan sebuah penelitian yang dapat melakukan klasifikasi data penderita penyakit stroke untuk mengidentifikasi pola-pola umum yang terjadi pada penderita dengan harapan dapat menurunkan angka penderita

¹ WHO, 2020: “*World Stroke Day*” <https://www.who.int/southeastasia/news/detail/29-10-2020/world-stroke-day-ms> diakses tanggal : 11 Agustus 2023”

² Situs web “Kementerian Kesehatan Republik Indonesia, 2022”. *Tingkatan Kualitas dan Layanan Stroke Lewat Transformasi Kesehatan*. Diakses pada 11 Agustus 2023

penyakit stroke di Indonesia. Dengan memahami faktor risiko tersebut, diharapkan dapat mengidentifikasi faktor-faktor risiko yang berkaitan dengan penyakit stroke, sehingga langkah-langkah pencegahan dapat diarahkan dengan lebih tepat dan efektif.

Berdasarkan permasalahan tersebut, pendekatan *data mining* dengan metode klasifikasi menjadi pilihan yang tepat. Klasifikasi adalah sebuah proses pengelompokkan kategori atau kelas yang didasari kesamaan pola ataupun sifat yang terdapat pada setiap data (Amini dkk., 2022). Beberapa algoritma klasifikasi yang umum digunakan adalah *Random Forest*, *Naïve Bayes*, *Decision Tree*, *K-Nearest Neighbor*, *Artificial Neural Network*, dan *Fuzzy*.

Penelitian sebelumnya (Azhar dkk., 2022), dilakukan perbandingan hasil klasifikasi data penyakit stroke menggunakan beberapa algoritma klasifikasi *data mining*, termasuk *Random Forest*, *Naïve Bayes*, dan *K-Nearest Neighbor*. Hasil penelitian tersebut menunjukkan bahwa algoritma *Random Forest* memiliki akurasi tertinggi, mencapai 72.1% untuk data yang seimbang dan 98.63% untuk data yang tidak seimbang. Berdasarkan penelitian tersebut menunjukkan bahwa Algoritma *Random Forest* memiliki kemampuan klasifikasi yang baik dan juga dipengaruhi oleh keseimbangan data. Hal ini terjadi karena hasil *Random Forest* tidak hanya menggunakan satu pohon keputusan melainkan menggunakan banyak pohon, sehingga dapat meningkatkan ketepatan klasifikasi (Fadlilah dkk., 2019).

Namun, dalam penggunaannya, algoritma *Random Forest* juga memiliki kelemahan, seperti memerlukan waktu yang cukup lama untuk memproses data karena melibatkan penggabungan sejumlah besar pohon keputusan untuk

menentukan kelas (Breiman, 2001). Untuk mengatasi hal ini, penggunaan teknik seleksi fitur menjadi solusi yang sangat efektif. Dengan menggunakan teknik ini, atribut-atribut yang paling relevan dapat diidentifikasi, sehingga proses pembuatan pohon keputusan menjadi lebih efisien dan akurat.

Algoritma Genetika adalah salah satu metode optimasi yang sering digunakan untuk seleksi fitur dalam pembelajaran mesin. Algoritma ini mengadopsi prinsip-prinsip seleksi alam dan genetika alami dan telah berhasil diterapkan dalam bidang *machine learning*. Dalam penelitian yang dilakukan oleh Busono (2020), Algoritma Genetika digunakan sebagai metode seleksi fitur pada algoritma *Naïve Bayes* untuk memprediksi performa siswa. Hasil penelitian tersebut menunjukkan bahwa akurasi *Naïve Bayes* tanpa optimasi adalah 91,15%. Namun, setelah optimasi menggunakan Algoritma Genetika, akurasi meningkat signifikan menjadi 97,21%.

Berdasarkan uraian latar belakang di atas, algoritma *Random Forest* dan Algoritma Genetika akan digunakan dalam klasifikasi data penderita penyakit stroke. Hasil dari penelitian ini diharapkan dapat membantu para medis dalam mengklasifikasi data penderita penyakit stroke dan memberikan edukasi kepada masyarakat untuk melakukan langkah-langkah pencegahan yang tepat.

1.3 Rumusan Masalah

Rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana perbandingan tingkat akurasi antara *Random Forest* dan *Random Forest* yang telah dioptimasi menggunakan Algoritma Genetika?

1.4 Tujuan Penelitian

Adapun berikut merupakan tujuan penelitian ini:

1. Menghasilkan sistem pengklasifikasian penyakit stroke menggunakan algoritma *Random Forest*.
2. Menganalisis perbandingan akurasi antara hasil dari algoritma *Random Forest* dengan hasil akurasi dari *Random Forest* yang telah dioptimasi menggunakan Algoritma Genetika.

1.5 Manfaat Penelitian

Berikut merupakan manfaat penelitian topik tugas akhir ini, diantaranya:

1. Sistem yang dihasilkan dapat digunakan untuk mengklasifikasikan data penderita penyakit stroke.
2. Penelitian ini diharapkan dapat menjadi referensi untuk penelitian selanjutnya dalam bidang yang sama.

1.6 Batasan Masalah

Batasan masalah dalam penelitian ini adalah dataset yang digunakan mempunyai 2 kategori. Kategori dengan nilai 0 untuk tidak stroke dan kategori dengan nilai 1 untuk stroke.

1.7 Sistematika Penulisan

Penelitian ini mengikuti struktur penulisan yang telah ditetapkan oleh Fakultas Ilmu Komputer Universitas Sriwijaya, diantaranya:

BAB I. PENDAHULUAN

Bab 1 menjelaskan tentang latar belakang penelitian, rumusan masalah penelitian, tujuan dan manfaat pada penelitian, batasan masalah penelitian, dan sistematika penulisan.

BAB II. KAJIAN LITERATUR

Bab ini membahas teori-teori dasar yang menjadi landasan penelitian, termasuk definisi stroke, pengklasifikasian data, algoritma yang dipakai, serta penelitian-penelitian sebelumnya yang relevan dengan topik penelitian.

BAB III. METODOLOGI PENELITIAN

Bab ini menguraikan tahapan yang akan dilaksanakan dalam penelitian. Setiap tahap perencanaan penelitian akan dijabarkan secara terperinci berdasarkan suatu kerangka kerja yang telah ditetapkan. Selain itu, bab ini juga mencakup perancangan manajemen proyek untuk pelaksanaan penelitian, yang bertujuan untuk kelancaran dan keberhasilan seluruh proses.

BAB IV. PENGEMBANGAN PERANGKAT LUNAK

Bab ini memaparkan seluruh proses pengembangan perangkat lunak dengan menerapkan metode *Rational Unified Process* (RUP).

BAB V. HASIL DAN ANALISIS PENELITIAN

Bab ini membahas hasil analisis dan perancangan yang telah dilakukan sebelumnya. Implementasi ini mencakup penerapan kesimpulan dari penelitian, termasuk pengujian perangkat lunak dan uji coba pada data penelitian untuk memvalidasi hasil yang diperoleh.

BAB VI. KESIMPULAN DAN SARAN

Bab ini memuat kesimpulan dari penjelasan yang telah disampaikan sebelumnya serta saran-saran yang diharapkan dapat memberikan manfaat untuk pengembangan penelitian lebih lanjut.

1.8 Kesimpulan

Dari penjelasan sebelumnya, dapat ditarik kesimpulan bahwa fokus penelitian adalah tentang cara mengklasifikasikan stroke menggunakan algoritma *Random Forest*, yang kemudian ditingkatkan akurasi melalui optimasi menggunakan Algoritma Genetika.

DAFTAR PUSTAKA

- akbar, F., Saputra, H. W., Maulaya, A. K., Hidayat, M. F., & Rahmadden, R. (2022). Implementasi Algoritma Decision Tree C4. 5 Dan Support Vector Regression Untuk Prediksi Penyakit Stroke: Implementation Of Decision Tree Algorithm C4. 5 And Support Vector Regression For Stroke Disease Prediction. *Malcom: Indonesian Journal Of Machine Learning And Computer Science*, 2(2), 61–67.
- Amanda, V., Rini, D. P., & Utami, A. S. (2023). *Optimasi Algoritma Random Forest Menggunakan Algoritma Genetika Dalam Proses Klasifikasi Penyakit Ginjal Kronis* [Phd Thesis]. Sriwijaya University.
- Amila, A., Sinaga, J., & Sembiring, E. (2018). Pencegahan Stroke Berulang Melalui Pemberdayaan Keluarga Dan Modifikasi Gaya Hidup. *Jurnal Abdimas*, 22(2), 143–150.
- Amini, N., Saragih, T. H., Faisal, M. R., Farmadi, A., & Abadi, F. (2022). Implementasi Algoritma Genetika Untuk Seleksi Fitur Pada Klasifikasi Genre Musik Menggunakan Metode Random Forest. *Jurnal Informatika Polinema*, 9(1), Article 1. <https://doi.org/10.33795/Jip.V9i1.1028>
- Aprillia, R. T. (T.T.). *Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya 2023*.
- Azhar, Y., Firdausy, A. K., & Amelia, P. J. (2022). Perbandingan Algoritma Klasifikasi Data Mining Untuk Prediksi Penyakit Stroke. *Sintech (Science And Information Technology) Journal*, 5(2), Article 2. <https://doi.org/10.31598/Sintechjournal.V5i2.1222>
- Binarwati, L., Mukhlash, I., & Soetrisno, S. (2017). Implementasi Algoritma Genetika Untuk Optimalisasi Random Forest Dalam Proses Klasifikasi Penerimaan Tenaga Kerja Baru: Studi Kasus Pt.Xyz. *Jurnal Sains Dan Seni Its*, 6(2), A78–A83. <https://doi.org/10.12962/J23373520.V6i2.26887>
- Biswas, N., Uddin, K. M. M., Rikta, S. T., & Dey, S. K. (2022). A Comparative Analysis Of Machine Learning Classifiers For Stroke Prediction: A

- Predictive Analytics Approach. *Healthcare Analytics*, 2, 100116.
<https://doi.org/10.1016/j.health.2022.100116>
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), Article 1.
<https://doi.org/10.1023/A:1010933404324>
- Busono, S. (2020). Optimasi Naive Bayes Menggunakan Algoritma Genetika Sebagai Seleksi Fitur Untuk Memprediksi Performa Siswa. *Jurnal Ilmiah Teknologi Informasi Asia*, 14(1), Article 1.
- Candra, E. N., Cholissodin, I., & Wihandika, R. C. (T.T.). Klasifikasi Status Gizi Balita Menggunakan Metode Optimasi Random Forest Dengan Algoritma Genetika (Studi Kasus: Puskesmas Cakru). *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer E-Issn*, 2548, 964x.
- Fadlilah, M. S., Wihandika, R. C., & Rahayudi, B. (T.T.). Klasifikasi Penurunan Fungsi Kognitif Pasien Stroke Menggunakan Metode Klasifikasi Random Forest.
- Gani, G., Haryeti, P., & Sopiah, P. (2023). Hubungan Modifikasi Gaya Hidup Dengan Pencegahan Stroke Ulang Pada Pasien Pasca Stroke Di Wilayah Kerja Puskesmas Tanjungsari. *Prepotif: Jurnal Kesehatan Masyarakat*, 7(1), 814–821.
- Hassanat, A., Almohammadi, K., Alkafaween, E., Abunawas, E., Hammouri, A., & Prasath, V. B. S. (2019). Choosing Mutation And Crossover Ratios For Genetic Algorithms—A Review With A New Dynamic Approach. *Information*, 10(12), 390. <https://doi.org/10.3390/info10120390>
- Kamila, S. A., Sulistijowati, R. S., & Susanto, I. (2023). Classification Of Heart Disease Using Decision Tree And Random Forest. *Seminar Nasional Teknologi & Sains*, 2(1), 7–12.
<https://proceeding.unpkediri.ac.id/index.php/stains/article/view/2816>
- Krishnaiah, V., Narsimha, D. G., & Chandra, D. N. S. (2014). Survey Of Classification Techniques In Data Mining. *International Journal Of Computer Sciences And Engineering*, 2.

- Malhotra, R., Singh, N., & Singh, Y. (2011). Genetic Algorithms: Concepts, Design For Optimization Of Process Controllers. *Computer And Information Science*, 4(2), 39.
- Nugroho, A. (2022). Analisa Splitting Criteria Pada Decision Tree Dan Random Forest Untuk Klasifikasi Evaluasi Kendaraan. *Jsitik: Jurnal Sistem Informasi Dan Teknologi Informasi Komputer*, 1(1), 41–49.
- Pambudi, R. E., Sriyanto, S., & Firmansyah, F. (2022). Klasifikasi Penyakit Stroke Menggunakan Algoritma Decision Tree C. 45. *Teknika*, 16(2), 221–226.
- Suryanegara, G. A. B., & Purbolaksono, M. D. (2021). Peningkatan Hasil Klasifikasi Pada Algoritma Random Forest Untuk Deteksi Pasien Penderita Diabetes Menggunakan Metode Normalisasi. *Jurnal Resti (Rekayasa Sistem Dan Teknologi Informasi)*, 5(1), 114–122.
- Susilawati, F., & Hk, N. (2018). *Faktor Resiko Kejadian Stroke Di Rumah Sakit* (1). 1, Article 1.
- Tia, T. K., & Nuryasin, I. (T.T.). *Model Simulasi Rational Unified Process Pada Pengembangan Perangkat Lunak* (4). 2(4), Article 4.
- Togatorop, P. R., Sianturi, M., Simamora, D., & Silaen, D. (2022). Optimizing Random Forest Using Genetic Algorithm For Heart Disease Classification. *Lontar Komputer : Jurnal Ilmiah Teknologi Informasi*, 13(1), 60. <https://doi.org/10.24843/Lkjiti.2022.V13.I01.P06>
- Widodo, A. M., Anwar, N., Irawan, B., Meria, L., & Wisnujati, A. (T.T.). *Komparasi Performansi Algoritma Pengklasifikasi Knn, Bagging Dan Random Forest Untuk Prediksi Kanker Payudara*.