

OPTIMASI ALGORITMA *NAÏVE BAYES* MENGGUNAKAN
ANT COLONY OPTIMIZATION UNTUK KLASIFIKASI DATA
PENDERITA PENYAKIT STROKE

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



Oleh :

Kurnia Minari

NIM : 09021182025004

Jurusan Teknik Informatika
FAKULTAS ILMU KOMPUTER UNIVERSITAS SRIWIJAYA

2024

LEMBAR PENGESAHAN SKRIPSI

**OPTIMASI ALGORITMA *NAIVE BAYES* MENGGUNAKAN
ANT COLONY OPTIMIZATION UNTUK KLASIFIKASI DATA
PENDERITA PENYAKIT STROKE**

Oleh :

KURNIA MINARI
NIM : 09021182025004

Palembang, 4 Januari 2024

Mengetahui,
Ketua Jurusan Teknik Informatika



Alvi Syahrini Utami, M.Kom.
NIP 197812222006042003

Pembimbing



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

TANDA LULUS UJIAN KOMPREHENSIF SKRIPSI

Pada hari Jum'at tanggal 22 Desember 2023 telah dilaksanakan ujian komprehensif skripsi oleh Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya

Nama : Kurnia Minari
NIM : 09021182025004
Judul : Optimasi Algoritma *Naïve Bayes* Menggunakan *Ant Colony Optimization* Untuk Klasifikasi Data Penderita Penyakit Stroke

dan dinyatakan LULUS.

1. Ketua Penguji

Rizki Kurniati, M.T.
NIP 199107122019032016



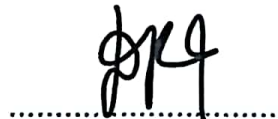
2. Penguji I

Annisa Darmawahyuni, M.Kom.
NIP 199006302023212044



3. Pembimbing I

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



Mengetahui,
Ketua Jurusan Teknik Informatika



Alvi Syahrini Utami, M.Kom
NIP. 197812222006042003

HALAMAN PERNYATAAN

Yang bertanda tangan dibawah ini:

Nama : Kurnia Minari
NIM : 09021182025004
Program Studi : Teknik Informatika
Judul : Optimasi Algoritma *Naïve Bayes* Menggunakan *Ant Colony Optimization* Untuk Klasifikasi Data Penderita Penyakit Stroke

Hasil pengecekan iThenticate/Turnitin: 13%

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 sebenarnya dan tidak ada paksaan oleh siapapun.

Palembang, Desember 2023



Kurnia Minari
09021182025004

MOTTO DAN PERSEMBAHAN

“Jangan kamu terburu-buru untuk sampai ke depan, karena kamu akan kehilangan banyak hal-hal penting.”

- D.O. Kyungsoo -

Kupersembahkan karya tulis ini kepada :

- Allah SWT
- Ibu dan Bapak
- Adik-adikku
- Keluarga besarku
- Sahabatku dan teman-teman
seperjuangan TI REG C 2020
- Dosen Pembimbing
- Almamater

ABSTRACT

Stroke is a dangerous disease that is included in the 10 deadliest diseases in Indonesia, with a very high mortality rate. To detect strokes, CT scans and MRI examinations are usually used. However, this examination is expensive and is often only carried out after someone has had a stroke. As a result, medical treatment can be delayed and the patient's condition can worsen. Therefore, it is very important for a method that can identify a patient's condition at an early stage to detect the risk of stroke by classifying whether the patient is experiencing a stroke or not. One method that is often used for classification is Naïve Bayes. However, this method has the weakness of passing data into certain classes even though the data is not relevant or relevant so it needs to be optimized by feature selection. From this, optimization was carried out using the Ant Colony Optimization algorithm for feature selection for classification using Naïve Bayes. In this research, feature selection will be carried out on the data that will be classified with Naïve Bayes using the Ant Colony Optimization algorithm. The accuracy of the Naïve Bayes method before and after being optimized with the Ant Colony Optimization algorithm was 86.45% and 95.84% without the use of resampling techniques, while with the application of resampling techniques, the accuracy was 78.69% and 79.91%.

Keywords: *Ant Colony Optimization, Classification, Feature Selection, Naïve Bayes, Stroke, Supervised Learning*

ABSTRAK

Stroke adalah penyakit berbahaya yang termasuk dalam 10 penyakit paling mematikan di Indonesia, dengan tingkat kematian yang sangat tinggi. Untuk mendeteksi stroke, biasanya digunakan pemeriksaan CT Scan dan MRI. Namun, pemeriksaan ini mahal dan sering kali baru dilakukan setelah seseorang terkena stroke. Akibatnya, penanganan medis bisa terlambat dan kondisi pasien bisa memburuk. Oleh karena itu, sangat penting untuk metode yang dapat mengidentifikasi kondisi pasien pada tahap awal guna mendeteksi risiko stroke dengan cara mengklasifikasikan apakah pasien tersebut mengalami stroke atau tidak. Salah satu metode yang sering digunakan untuk klasifikasi adalah *Naïve Bayes*. Namun, metode ini memiliki kelemahan dengan meloloskan data kedalam kelas tertentu meskipun data tersebut tidak relevan atau bersangkutan sehingga perlu dioptimasi dengan seleksi fitur. Dari hal tersebut dilakukan optimasi menggunakan algoritma *Ant Colony Optimization* untuk seleksi fitur pada klasifikasi menggunakan *Naïve Bayes*. Pada penelitian ini akan dilakukan seleksi fitur pada data yang akan diklasifikasikan dengan *Naïve Bayes* menggunakan algoritma *Ant Colony Optimization*. Akurasi metode *Naïve Bayes* sebelum dan setelah dioptimalkan dengan algoritma *Ant Colony Optimization* adalah 86,45% dan 95,84% tanpa penggunaan teknik *resampling*, sementara dengan penerapan teknik *resampling*, akurasinya menjadi 78,69% dan 79,91%.

Kata Kunci: *Ant Colony Optimization, Feature Selection, Klasifikasi, Naïve Bayes, Stroke, Supervised Learning*

KATA PENGANTAR

Puji syukur kepada Allah SWT atas berkat dan rahmat-Nya yang telah diberikan kepada Penulis sehingga dapat menyelesaikan Tugas Akhir ini dengan baik. Tugas akhir ini disusun untuk memenuhi salah satu syarat guna menyelesaikan pendidikan program Strata-1 pada Fakultas Ilmu Komputer Program Studi Teknik Informatika di Universitas Sriwijaya.

Penyusunan laporan tugas akhir ini melibatkan bimbingan, bantuan, dukungan, dan arahan dari berbagai pihak. Oleh karena itu, penulis ingin menyampaikan rasa terima kasih kepada semua yang telah memberikan kontribusi dalam menyelesaikan laporan tugas akhir ini, terutama kepada:

1. Allah SWT atas segala rahmat, petunjuk, dan karunia-Nya.
2. Nabi Muhammad SAW atas petunjuk dan keberkahan yang terinspirasi dari ajaran dan teladan beliau dalam setiap langkah perjalanan penyusunan skripsi ini.
3. Ibu, Bapak, adik-adik dan seluruh keluarga besarku yang selalu medoakan serta memberikan dukungan baik secara moril maupun materil.
4. Prof. DR. Erwin, S.Si., M.Si. selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
5. Alvi Syahrini Utami, M.Kom. selaku Ketua Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya dan Mastura Diana Marieska, M.T. selaku Sekretaris Jurusan Teknik Informatika.

6. Dian Palupi Rini, M.Kom., Ph.D. selaku Dosen Pembimbing I yang telah menyediakan waktu, senantiasa memberikan bimbingan, masukan, dan motivasi untuk menyelesaikan tugas akhir ini dengan baik.
7. Rizki Kurniati, M.T. selaku Ketua Penguji tugas akhir yang telah memberikan saran dan masukan.
8. Annisa Darmawahyuni, M.Kom. selaku Dosen Penguji I tugas akhir yang telah memberikan saran dan masukan yang membangun.
9. Julian Supardi, M.T., Ph.D. selaku Dosen Pembimbing Akademik yang senantiasa memberikan bimbingan selama masa perkuliahan.
10. Seluruh Dosen Jurusan Teknik Informatika dan juga Dosen Fakultas Ilmu Komputer yang telah banyak memberikan ilmu selama masa perkuliahan.
11. Seluruh staf administrasi dan pegawai Fakultas Ilmu Komputer yang telah membantu dalam urusan administrasi.
12. Teman-teman seperjuanganku, Aini, Koko, Dini, Citra, Melsra, Betha, Tiara, Meri dan teman-teman TI REG 2020 yang telah menanggung beratnya kehidupan layo dengan cerita yang menarik selama menjalani perkuliahan dan pengerjaan tugas akhir serta untuk selalu berjuang bersama dalam menempuh ilmu.
13. Saudara-saudaraku semasa di BEM KM Fasilkom UNSRI 2021, Zhafira, Sheva, Kak Rifki dan kaum nokturnal atas rasa solidaritas, bantuan, dukungan, saran dan menjadi tempat nyaman yang telah mewarnai dunia perkuliahan.

14. Sahabat-sahabat mainku Carissa, Sabina, Febi dan Widia atas *support system* disaat suka maupun duka, selalu mendoakan kelancaran, selalu menemani Penulis dan menjadi tempat ternyaman berbagi keluh kesah dalam menyelesaikan studi serta pengerjaan tugas akhir.
15. Orang-orang terdekat yang kusayangi dan menyayangiku yang selalu memberikan motivasi, mendoakan kelancaran dan menjadi tempat curhat dalam masa perkuliahan dan penulisan dan mencari solusi atas semua permasalahan selama penulisan.
16. Seluruh pihak yang tidak dapat disebutkan satu persatu yang telah banyak membantu memberikan pemikiran demi kelancaran dalam pembuatan dan pengerjaan tugas akhir ini.

Penulis berharap semoga tulisan ini dapat menjadi bahan pembelajaran di masa yang akan datang. Penulis juga menyadari dalam penyusunan Tugas Akhir ini masih terdapat banyak kekurangan disebabkan keterbatasan pengetahuan dan pengalaman, oleh karena itu kritik dan saran yang membangun sangat diharapkan untuk kemajuan penelitian selanjutnya. Akhir kata semoga Tugas Akhir ini dapat berguna dan bermanfaat bagi kita semua.

Palembang, Desember 2023

Penulis

DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN PENGESAHAN SKRIPSI.....	ii
TANDA LULUS UJIAN SIDANG SKRIPSI.....	iii
HALAMAN PERNYATAAN	iv
MOTTO DAN PERSEMBAHAN	v
ABSTRACT.....	vi
ABSTRAK.....	vii
KATA PENGANTAR.....	viii
DAFTAR ISI.....	xi
DAFTAR TABEL.....	xv
DAFTAR GAMBAR	xvii
DAFTAR LAMPIRAN	xviii
BAB I PENDAHULUAN.....	I-1
1.1 Pendahuluan	I-1
1.2 Latar Belakang.....	I-1
1.3 Rumusan Masalah	I-5
1.4 Tujuan Penelitian.....	I-5
1.5 Manfaat Penelitian.....	I-5
1.6 Batasan Masalah.....	I-6
1.7 Sistematika Penulisan.....	I-6
1.8 Kesimpulan.....	I-8
BAB II KAJIAN LITERATUR.....	II-1
2.1 Pendahuluan	II-1
2.2 Landasan Teori	II-1
2.2.1 <i>Supervised Learning</i>	II-1
2.2.2 Klasifikasi	II-2
2.2.3 Algoritma <i>Naïve Bayes</i>	II-2
2.2.4 <i>Feature Selection</i>	II-6
2.2.5 Algoritma <i>Ant Colony Optimization</i>	II-7
2.2.6 Stroke	II-11

2.3	Penelitian Lain Yang Relevan.....	II-12
2.4	Kesimpulan.....	II-13
BAB III METODOLOGI PENELITIAN.....		III-1
3.1	Pendahuluan	III-1
3.2	Pengumpulan Data.....	III-1
3.2.1	Jenis Data.....	III-1
3.2.2	Sumber Data	III-1
3.2.3	Metode Pengumpulan Data.....	III-2
3.3	Tahapan Penelitian.....	III-2
3.3.1	Kerangka Kerja	III-4
3.3.2	Kriteria Pengujian	III-6
3.3.3	Format Data Pengujian	III-7
3.3.4	Alat yang Digunakan dalam Pelaksanaan Penelitian.....	III-8
3.3.5	Pengujian Penelitian	III-8
3.3.6	Analisis Hasil Pengujian dan Membuat Kesimpulan	III-9
3.4	Metode Pengembangan Perangkat Lunak	III-9
3.4.1	Fase Insepsi.....	III-10
3.4.2	Fase Elaborasi	III-10
3.4.3	Fase Konstruksi.....	III-11
3.4.4	Fase Transisi	III-11
3.5	Manajemen Proyek Penelitian.....	III-11
3.6	Kesimpulan.....	III-17
BAB IV PENGEMBANGAN PERANGKAT LUNAK		IV-1
4.1	Pendahuluan	IV-1
4.2	<i>Rational Unified Process (RUP)</i>	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-4
4.2.1.3.1.	Analisis Kebutuhan Perangkat Lunak	IV-4
4.2.1.3.2.	Analisis Data	IV-4
4.2.1.3.3.	Analisis Seleksi Fitur dengan <i>Ant Colony Optimization</i>	IV-5

4.2.1.3.4. Analisis Klasifikasi dengan <i>Naïve Bayes</i>	IV-11
4.2.1.3.5. Perancangan Perangkat Lunak	IV-16
4.2.2. Fase Elaborasi	IV-19
4.2.2.1. Pemodelan Bisnis.....	IV-20
4.2.2.1.1. Perancangan Antarmuka.....	IV-20
4.2.2.1.2. Perancangan Data	IV-21
4.2.2.2. Kebutuhan Sistem.....	IV-22
4.2.2.3. Diagram	IV-22
4.2.2.3.1. <i>Activity</i> Diagram.....	IV-22
4.2.2.3.2. <i>Sequence</i> Diagram.....	IV-24
4.2.3. Fase Kontruksi	IV-25
4.2.3.1. Pemodelan Bisnis.....	IV-25
4.2.3.2. Diagram Kelas	IV-26
4.2.3.3. Implementasi.....	IV-26
4.2.3.3.1. Implementasi Kelas	IV-26
4.2.3.3.2. Implementasi Antarmuka	IV-27
4.2.4. Fase Transisi	IV-28
4.2.4.1. Pemodelan Bisnis.....	IV-28
4.2.4.2. Kebutuhan Sistem.....	IV-29
4.2.4.3. Rencana Pengujian.....	IV-29
4.2.4.3.1. Rencana Pengujian <i>Use Case</i> Klasifikasi dengan <i>Naïve Bayes</i>	IV-29
4.2.4.3.2. Rencana Pengujian <i>Use Case</i> Melakukan Seleksi Fitur dengan ACO dan Klasifikasi dengan <i>Naïve Bayes</i>	IV-30
4.2.4.3.3. Implementasi	IV-31
4.2.4.3.4. Pengujian <i>Use Case</i> Klasifikasi dengan <i>Naïve Bayes</i>	IV-31
4.2.4.3.5. Pengujian <i>Use Case</i> Melakukan Seleksi Fitur dengan ACO dan Klasifikasi dengan <i>Naïve Bayes</i>	IV-32
4.3 Kesimpulan.....	IV-33
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.	Data Hasil Pengujian Seleksi Fitur dengan <i>Ant Colony Optimization</i>	V-3
5.2.3.	Data Hasil Pengujian Menggunakan <i>Naïve Bayes</i>	V-6
5.2.4.	Data Hasil Pengujian Klasifikasi Menggunakan <i>Naïve Bayes</i>	V-8
5.2.5.	Data Hasil Pengujian Klasifikasi dengan <i>Naïve Bayes</i> dan ACO.	V-11
5.3.	Analisis Hasil Penelitian.....	V-14
5.3.1.	Hasil Perbandingan Klasifikasi Menggunakan <i>Naïve Bayes</i> dengan dan tanpa Seleksi Fitur Menggunakan ACO	V-14
5.4.	Kesimpulan.....	V-15
BAB VI KESIMPULAN DAN SARAN		VI-1
6.1.	Pendahuluan	VI-1
6.2.	Kesimpulan.....	VI-1
6.3.	Saran.....	VI-2
DAFTAR PUSTAKA		xix
LAMPIRAN		xxiv

DAFTAR TABEL

Tabel III-1	Rancangan Tabel Hasil Pengujian Seleksi Fitur dengan <i>Ant Colony Optimization</i>	III-7
Tabel III-2	Rancangan Tabel Hasil Pengujian dengan <i>Naïve Bayes</i>	III-7
Tabel III-3	Rancangan Tabel Pengujian Klasifikasi dengan <i>Naïve Bayes</i>	III-7
Tabel III-4	Rancangan Tabel Hasil Pengujian Klasifikasi dengan <i>Naïve Bayes</i> dan ACO	III-8
Tabel III-5	Rancangan Tabel Hasil Perbandingan Klasifikasi dengan dan Tanpa Seleksi Fitur	III-9
Tabel III- 6.	Tabel Penjadwalan Penelitian dalam Bentuk <i>Work Breakdown Structure (WBS)</i>	III-12
Tabel IV- 1.	Kebutuhan Fungsional Perangkat Lunak	IV-3
Tabel IV- 2.	Kebutuhan Non-Fungsional Perangkat Lunak	IV-3
Tabel IV- 3.	Matriks Jarak antar Node Kota.....	IV-5
Tabel IV- 4.	Hasil Perhitungan ACO pada Jarak antar Kota	IV-9
Tabel IV- 5.	Data <i>Training</i>	IV-11
Tabel IV- 6.	Data <i>Testing</i>	IV-12
Tabel IV- 7.	Perhitungan Nilai Probabilitas pada Data <i>Testing</i>	IV-12
Tabel IV- 8.	Tabel Fakta	IV-16
Tabel IV- 9.	Definisi Pengguna <i>Use Case</i>	IV-17
Tabel IV- 10.	Definisi Diagram <i>Use Case</i>	IV-17
Tabel IV- 11.	Skenario Melakukan Klasifikasi <i>Naïve Bayes</i>	IV-17
Tabel IV- 12.	Skenario Seleksi Fitur dengan ACO dan Klasifikasi dengan <i>Naïve Bayes</i>	IV-18
Tabel IV- 13.	Implementasi Kelas	IV-27
Tabel IV- 14.	Rencana Pengujian <i>Use Case</i> Klasifikasi dengan <i>Naïve Bayes</i>	IV-29
Tabel IV- 15.	Rencana Pengujian <i>Use Case</i> Melakukan Seleksi Fitur dengan ACO dan Klasifikasi dengan <i>Naïve Bayes</i>	IV-30
Tabel IV- 16.	Pengujian <i>Use Case</i> Klasifikasi dengan <i>Naïve Bayes</i>	IV-31

Tabel IV- 17. Pengujian <i>Use Case</i> Melakukan Seleksi Fitur dengan ACO dan Klasifikasi dengan <i>Naïve Bayes</i>	IV-32
Tabel V- 1. Hasil Pengujian Seleksi Fitur dengan <i>Ant Colony Optimization</i>	V-3
Tabel V- 2. Hasil Pengujian Seleksi Fitur dengan <i>Ant Colony Optimization</i> Menggunakan Teknik <i>Resampling</i>	V-4
Tabel V- 3. Hasil Pengujian Menggunakan <i>Naïve Bayes</i>	V-6
Tabel V- 4. Hasil Pengujian Menggunakan <i>Naïve Bayes</i> dengan Teknik <i>Resampling</i>	V-7
Tabel V- 5. Hasil Pengujian Klasifikasi Menggunakan <i>Naïve Bayes</i>	V-8
Tabel V- 6. <i>Confusion Matrix</i> Klasifikasi Menggunakan <i>Naïve Bayes</i> tanpa Teknik <i>Resampling</i>	V-9
Tabel V- 7. Hasil Klasifikasi Menggunakan <i>Naïve Bayes</i> tanpa Teknik <i>Resampling</i>	V-9
Tabel V- 8. <i>Confusion Matrix</i> Klasifikasi Menggunakan <i>Naïve Bayes</i> dengan Teknik <i>Resampling</i>	V-10
Tabel V- 9. Hasil Klasifikasi Menggunakan <i>Naïve Bayes</i> dengan Teknik <i>Resampling</i>	V-10
Tabel V- 10. Hasil Pengujian Klasifikasi dengan <i>Naïve Bayes</i> dan ACO	V-11
Tabel V- 11. <i>Confusion Matrix</i> Klasifikasi Menggunakan <i>Naïve Bayes</i> dan ACO tanpa Teknik <i>Resampling</i>	V-12
Tabel V- 12. Hasil Klasifikasi Menggunakan <i>Naïve Bayes</i> dan ACO tanpa Teknik <i>Resampling</i>	V-12
Tabel V- 13. <i>Confusion Matrix</i> Klasifikasi Menggunakan <i>Naïve Bayes</i> dan ACO dengan Teknik <i>Resampling</i>	V-13
Tabel V- 14. Hasil Klasifikasi Menggunakan <i>Naïve Bayes</i> dan ACO dengan Teknik <i>Resampling</i>	V-13
Tabel V- 15. Hasil Perbandingan Klasifikasi Menggunakan <i>Naïve Bayes</i> dengan dan tanpa Seleksi Fitur Menggunakan ACO	V-14

DAFTAR GAMBAR

Gambar II-1. Tahapan Proses Klasifikasi <i>Naïve Bayes</i>	II-6
Gambar II-2. Tahapan Proses ACO	II-11
Gambar III- 1. Tahapan Penelitian	III-2
Gambar III- 2. Kerangka Kerja	III-4
Gambar III- 3. Tahap Pre-processing Data	III-5
Gambar IV- 1. Diagram Use Case	IV-16
Gambar IV- 2. Rancangan Antarmuka Perangkat Lunak Klasifikasi dengan <i>Naïve Bayes</i>	IV-20
Gambar IV- 3. Rancangan Antarmuka Perangkat Lunak Seleksi Fitur dengan ACO dan Klasifikasi dengan <i>Naïve Bayes</i>	IV-21
Gambar IV- 4. Diagram Aktivitas Klasifikasi dengan <i>Naïve Bayes</i>	IV-23
Gambar IV- 5. Diagram Aktivitas Seleksi Fitur dengan ACO dan Klasifikasi dengan <i>Naïve Bayes</i>	IV-23
Gambar IV- 6. Diagram <i>Sequence</i> Klasifikasi dengan <i>Naïve Bayes</i>	IV-24
Gambar IV- 7. Diagram <i>Sequence</i> Seleksi Fitur dengan ACO dan Klasifikasi dengan <i>Naïve Bayes</i>	IV-25
Gambar IV- 8. Diagram Kelas.....	IV-26
Gambar IV- 9. Antarmuka Klasifikasi dengan <i>Naïve Bayes</i>	IV-27
Gambar IV- 10. Antarmuka Klasifikasi dengan <i>Naïve Bayes</i> dan ACO	IV-28

DAFTAR LAMPIRAN

Lampiran 1. Dataset	xxiv
Lampiran 2. Kode Program	xxv

BAB I

PENDAHULUAN

1.1 Pendahuluan

Pada bab pendahuluan akan membahas tentang latar belakang masalah, rumusan masalah, tujuan dan manfaat dari penelitian, batasan masalah dan sistematika penulisan, serta berisikan penjelasan secara keseluruhan mengenai gambaran umum dari penelitian yang dilakukan.

1.2 Latar Belakang

Stroke atau *Cerebro Vaskuler Accident* (CVA) merupakan salah satu penyakit berbahaya dengan tingkat kematian yang sangat tinggi. Menurut data dari Organisasi Kesehatan Dunia atau *World Health Organization* (WHO) 2020, stroke menjadi penyakit dengan tingkat kematian tertinggi kedua didunia setelah penyakit jantung dan penyebab kecacatan nomor tiga di dunia¹⁾. Di Indonesia, stroke sendiri masuk kedalam 10 jenis penyakit yang paling mematikan. Kementerian Kesehatan Republik Indonesia tahun 2018 menyebutkan bahwa tingkat prevalensi penyakit stroke di Indonesia masih cukup tinggi yaitu 10.9 %. Dan disebutkan juga stroke menjadi penyebab kematian nomor satu disemua umur yaitu dengan presentase 15.4%²⁾.

Penyakit stroke didefinisikan sebagai kondisi ketika pasokan darah ke otak mengalami gangguan akibat sumbatan atau pecahnya pembuluh darah di otak yang

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⁾ Kemenkes RI, 2018 : “Potret Sehat Indonesia dari Riskesdas 2018” <http://www.depkes.go.id/article/view/18110200003/> diakses tanggal : 13 Agustus 2023.

menyebabkan kerusakan pada fungsi syaraf seperti kecacatan, kelumpuhan hingga dapat mengakibatkan kematian (Hutama et al., 2018). Stroke dapat terjadi pada siapa saja, tetapi risiko meningkatnya penyakit stroke dapat disebabkan oleh faktor-faktor seperti faktor kesehatan (hipertensi, kolesterol, obesitas dan jantung), faktor gaya hidup (merokok, kurang olahraga, mengonsumsi obat-obatan terlarang dan kecanduan alkohol), faktor keturunan, dan faktor usia (Haryadi et al., 2021).

Pendeteksian penyakit stroke dapat dilakukan dengan pemeriksaan menggunakan alat berupa *Computed Tomography (CT) Scan* (Pujiastuti & Azaria, 2018) dan *Magnet Resonance Imaging (MRI)* yang meliputi pemeriksaan fisik umum dan neurologis (Setiawan, 2021). Meskipun efektif, pemeriksaan dengan alat ini memiliki biaya yang tinggi, dan seringkali dilakukan setelah seseorang sudah terkena penyakit stroke, yang dapat mengakibatkan kondisi pasien memburuk akibat keterlambatan penanganan medis. Pemrosesan data citra, yang diperlukan dalam pemeriksaan ini, melibatkan ekstraksi fitur kompleks dan membutuhkan daya komputasi yang signifikan, serta waktu penyimpanan yang lama (Asri & Firmansyah, 2018).

Untuk meningkatkan efisiensi dalam mendeteksi tingkat awal risiko penyakit stroke, pendekatan menggunakan data tabular dari rekam medis pasien dapat menjadi solusi yang lebih efisien. Pemrosesan pada data tabular lebih mudah dan cepat karena strukturnya teratur, dengan waktu komputasi yang lebih rendah dibandingkan dengan data citra, dan membutuhkan lebih sedikit sumber daya memori (Badaro et al., 2023). Dengan demikian, penerapan sistem deteksi risiko

penyakit stroke menggunakan data tabular dapat menjadi solusi yang lebih efisien secara komputasional.

Sistem deteksi awal resiko penyakit stroke menggunakan data tabular pernah dikembangkan oleh beberapa peneliti yakni dengan judul Sistem Pakar Deteksi Dini Penyakit Stroke menggunakan metode *Naïve Bayes-Certainty Factor* dimana hasil penelitian ini menghasilkan akurasi sebesar 84% (Hutama et al., 2018). Kemudian penelitian selanjutnya yang dilakukan yakni Sistem Pendukung Keputusan Deteksi Dini pada Penyakit Stroke menggunakan metode Dempster *Shafer* dimana hasil penelitian ini menghasilkan akurasi sebesar 89% (Indraswari et al., 2015).

Salah satu teknik yang dimanfaatkan dalam pendeteksian awal resiko penyakit stroke agar dapat mengetahui apakah pasien tersebut menderita penyakit stroke atau tidak yaitu teknik klasifikasi. Klasifikasi adalah teknik yang digunakan untuk menempatkan suatu elemen ke dalam kategori yang telah ditetapkan sebelumnya, dimana elemen tersebut dimasukkan dalam kategori berdasarkan persamaan karakter data elemen tersebut (Paramitha et al., 2023). Terdapat berbagai jenis teknik pengklasifikasian salah satunya adalah metode *Naïve Bayes*.

Algoritma *Naïve Bayes* dalam klasifikasi sesuai untuk diimplementasikan pada data dengan jumlah yang banyak dan memiliki kemampuan untuk mengatasi data yang tidak lengkap (*missing value*), serta dapat menangani atribut yang beragam dan mengatasi gangguan yang mungkin muncul dalam data (Arifin & Ariesta, 2019). Algoritma ini memiliki kekurangan yang dapat mengurangi tingkat akurasinya karena algoritma ini tidak memiliki bobot atau semua bobot pada

atributnya bernilai sama sehingga tetap meloloskan semua atribut walaupun atribut itu tidak bersangkutan (Amalia, 2018). Untuk dapat meningkatkan akurasi pada algoritma *Naïve Bayes* perlu dilakukan optimasi dengan menghilangkan semua atribut yang tidak memadai melalui seleksi fitur dengan menggunakan Algoritma *Ant Colony Optimization* (Renuka et al., 2015).

Pada penelitian sebelumnya menunjukkan algoritma *Ant Colony Optimization* dapat meningkatkan performa suatu algoritma dengan melakukan seleksi fitur, penelitian yang dilakukan oleh (Renuka et al., 2015) menunjukkan bahwa algoritma *Naïve Bayes* mengalami peningkatan akurasi setelah dioptimasi dengan melakukan seleksi fitur menggunakan algoritma *Ant Colony Optimization* dari 77% menjadi 84%. Penelitian lain juga dilakukan oleh (Yusa et al., 2018) menunjukkan bahwa optimasi dengan seleksi fitur menggunakan algoritma *Ant Colony Optimization* dapat meningkatkan performa algoritma *K-Medoids* dari nilai akurasi *clustering* tertinggi sebesar 88.29% menjadi 97.65% sehingga algoritma ini digunakan penulis dalam penelitian karena dapat menemukan subset fitur yang optimal atau hampir optimal dengan waktu yang cukup efisien, terutama ketika data memiliki kompleksitas yang tinggi.

Berdasarkan uraian diatas, maka penulis akan meneliti pengaruh penggunaan *Ant Colony Optimization* dalam mengoptimasi algoritma *Naïve Bayes* untuk klasifikasi data penderita penyakit stroke. Dengan mengoptimasi algoritma *Naïve Bayes* menggunakan algoritma *Ant Colony Optimization*, diharapkan sistem dapat melakukan pendeteksian awal tingkat resiko penyakit stroke lebih efektif dan efisien.

1.3 Rumusan Masalah

Berdasarkan uraian latar belakang masalah diatas, maka dirumuskan masalah bagaimana *Ant Colony Optimization* dapat mengoptimasi algoritma *Naïve Bayes* pada klasifikasi penderita penyakit stroke, untuk melakukan penelitian ini akan dibagi dalam *research question*, sebagai berikut:

1. Bagaimana mekanisme algoritma *Naïve Bayes* untuk melakukan klasifikasi dan algoritma *Ant Colony Optimization* untuk melakukan seleksi fitur?
2. Apakah hasil akurasi dari klasifikasi penderita penyakit stroke dengan algoritma *Naïve Bayes* lebih optimal dengan menggunakan algoritma *Ant Colony Optimization*?

1.4 Tujuan Penelitian

Tujuan penelitian didasarkan pada rumusan masalah yang telah dijelaskan adalah sebagai berikut:

1. Memahami mekanisme algoritma *Naïve Bayes* dan algoritma *Ant Colony Optimization*.
2. Mengetahui apakah algoritma *Ant Colony Optimization* dapat meningkatkan hasil akurasi dari algoritma *Naïve Bayes* pada klasifikasi penderita penyakit stroke.

1.5 Manfaat Penelitian

Manfaat penelitian yang dapat diperoleh adalah sebagai berikut:

1. Dapat mengimplementasikan optimasi algoritma *Naïve Bayes* dengan algoritma *Ant Colony Optimization* pada sistem pengklasifikasi penderita penyakit stroke.
2. Sistem yang dibangun dapat membantu melakukan pendeteksian awal terhadap pasien penderita penyakit stroke.

1.6 Batasan Masalah

Batasan masalah yang akan dibahas pada penelitian adalah sebagai berikut:

1. Klasifikasi data penderita penyakit stroke tanpa mempertimbangkan jenis stroke.
2. Data uji yang digunakan dalam penelitian adalah *Stroke Prediction Dataset* yang diperoleh dari *Kaggle Dataset*³.

1.7 Sistematika Penulisan

Sistematika penulisan tugas akhir ini berisi penjelasan singkat bab perbab mengikuti standar penulisan tugas akhir Fakultas Ilmu Komputer Universitas Sriwijaya adalah sebagai berikut:

BAB I PENDAHULUAN

Pada bab ini berisi tentang pembahasan mengenai latar belakang, rumusan masalah, tujuan dan manfaat penelitian, batasan masalah atau ruang lingkup, metodologi penelitian, dan sistematika penulisan.

³) <https://www.kaggle.com/datasets/fedesoriano/stroke-prediction-dataset>

BAB II KAJIAN LITERATUR

Pada bab ini berisi tentang pembahasan mengenai keseluruhan dasar teori yang digunakan mulai dari definisi sistem, informasi mengenai domain yaitu stroke, dan semua yang digunakan dalam tahapan analisis, perancangan, dan implementasi.

BAB III METODOLOGI PENELITIAN

Pada bab ini berisi tentang pembahasan mengenai metode-metode yang akan digunakan pada penelitian. Setiap rencana dari metode penelitian dideskripsikan secara rinci berdasarkan kerangka kerja yang dibuat. Dilanjutkan dengan perancangan manajemen proyek penelitian.

BAB IV PENGEMBANGAN PERANGKAT LUNAK

Pada bab ini berisi tentang pembahasan mengenai perancangan dan lingkungan pelaksanaan, mencakup analisis masalah yang dihadapi dalam studi dan perancangan perangkat lunak untuk meningkatkan keakuratan algoritma *Naïve Bayes* dengan memanfaatkan *Ant Colony Optimization* dalam mengklasifikasikan penderita penyakit stroke. Perangkat lunak yang dirancang ini akan berfungsi sebagai alat penelitian dalam konteks penelitian ini.

BAB V HASIL DAN ANALISIS PENELITIAN

Pada bab ini berisi tentang pembahasan mengenai hasil analisis dan perancangan sebelumnya, implementasi ini melibatkan penerapan kesimpulan yang dihasilkan dari penelitian. Proses ini mencakup pengujian perangkat lunak dan uji coba pada data penelitian untuk memvalidasi hasil yang telah ditemukan.

BAB VI KESIMPULAN DAN SARAN

Pada bab ini berisi tentang kesimpulan dari penjelasan yang telah diberikan sebelumnya dan saran yang diharapkan dapat memberikan manfaat untuk pengembangan lebih lanjut dalam penelitian ini.

1.8 Kesimpulan

Berdasarkan masalah dari penelitian yang telah dijelaskan dilakukan penelitian terhadap data-data yang akan digunakan. Kemudian melakukan implementasi dengan data mining menggunakan algoritma *Naïve Bayes* yang dioptimasi dengan menggunakan *Ant Colony Optimization* untuk klasifikasi penderita penyakit stroke.

DAFTAR PUSTAKA

- Amalia, H. (2018). Perbandingan Metode Data Mining SVM dan NN Untuk Klasifikasi Penyakit Ginjal Kronis. *Jurnal PILAR Nusa Mandiri*, 14(1), 1–6.
- Annur, H. (2018). Klasifikasi Masyarakat Miskin Menggunakan Metode Naive Bayes. *ILKOM Jurnal Ilmiah*, 10(2), 160–165. <https://doi.org/10.33096/ilkom.v10i2.303.160-165>
- Arifin, T., & Ariesta, D. (2019). Prediksi Penyakit Ginjal Kronis Menggunakan Algoritma Naive Bayes Classifier Berbasis Particle Swarm Optimization. *Jurnal Tekno Insentif*, 13(1), 26–30. <https://doi.org/10.36787/jti.v13i1.97>
- Asri, J. S., & Firmansyah, G. (2018). Implementasi Objek Detection dan Tracking Menggunakan Deep Learning Untuk Pengolahan Citra Digital.
- Astuti, N. (2017). Hubungan Faktor Usia, Jenis Kelamin, Pekerjaan Dengan Jenis Stroke. *Jurnal Keperawatan Flora*, 10(2), 29–42.
- Badaro, G., Saeed, M., & Papotti, P. (2023). Transformers for Tabular Data Representation: A Survey of Models and Applications. *Transactions of the Association for Computational Linguistics*, 11, 227–249. https://doi.org/10.1162/tacl_a_00544
- Byna, A., & Basit, M. (2020). Penerapan Metode Adaboost Untuk Mengoptimasi Prediksi Penyakit Stroke Dengan Algoritma Naive Bayes. *Jurnal Sisfokom (Sistem Informasi dan Komputer)*, 9(3), 407–411. <https://doi.org/10.32736/sisfokom.v9i3.1023>

- Chen, B., Chen, L., & Chen, Y. (2013). Efficient ant colony optimization for image feature selection. *Special Issue on Machine Learning in Intelligent Image Processing*, 93(6), 1566–1576. <https://doi.org/10.1016/j.sigpro.2012.10.022>
- Haryadi, D., Atmaja, D. M. U., Hakim, A. R., & Suwaryo, N. (2021). Identifikasi Tingkat Resiko Penyakit Stroke Menggunakan Algoritma Regresi Linear Berganda. *SNTEM*, 1, 1198–1207. <https://doi.org/10.53026/sntem.v1i2.589>
- Hiran, K. K., Jain, R. K., Lakhwani, K., & Doshi, R. (2022). *Machine Learning: Master Supervised and Unsupervised Learning Algorithms with Real Examples* (1st ed.). BPB Publication, India. <https://books.google.com/books?hl=id&lr=&id=4VVDEAAAQBAJ&oi=fnd&pg=PT25&dq=related:Ro-Xp0Nra0YJ:scholar.google.com/&ots=OM8cDAK1kN&sig=qeKNXgygZFCloTMWIWum1Dvg87E>
- Hutama, R. S., Hidayat, N., & Santoso, E. (2018). Sistem Pakar Deteksi Dini Penyakit Stroke Menggunakan Metode Naïve Bayes-Certainty Factor. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 2(11), 4333–4339.
- Indraswari, D. P., Soebroto, A. A., & Marhaendraputro, E. A. (2015). Sistem Pendukung Keputusan Deteksi Dini Penyakit Stroke Menggunakan Metode Dempster-Shafer. *Journal of Enviromental Engineering and Sustainable Technology*, 2(2), 97–104. <https://doi.org/10.21776/ub.jeest.2015.002.02.6>

- Karjono, Moedjiono, & Kurniawan, D. (2016). Ant Colony Optimization. *Jurnal TICOM*, 4(3), 119–125.
- Mustafa, M. S., Ramadhan, M. R., & Thenata, A. P. (2018). Implementasi Data Mining untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier. *Creative Information Technology Journal*, 4(2), 151. <https://doi.org/10.24076/citec.2017v4i2.106>
- Nurdiawan, O., & Salim, N. (2018). Penerapan Data Mining Pada Penjualan Barang Menggunakan Metode Metode Naive Bayes Classifier Untuk Optimasi Strategi Pemasaran. *Jurnal Teknologi Informasi dan Komunikasi*, 84–95.
- Nurhayati, Busman, & Iswara, R. P. (2019). Pengembangan Algoritma Unsupervised Learning Technique Pada Big Data Analysis di Media Sosial sebagai media promosi Online Bagi Masyarakat. *JURNAL TEKNIK INFORMATIKA*, 12(1), 79–96. <https://doi.org/10.15408/jti.v12i1.11342>
- Nurlaelasari, E., Supriyadi, S., & Lenggana, U. T. (2018). Penerapan Algoritma Ant Colony Optimization Menentukan Nilai Optimal Dalam Memilih Objek Wisata Berbasis Android. *Simetris: Jurnal Teknik Mesin, Elektro dan Ilmu Komputer*, 9(1), 287–298. <https://doi.org/10.24176/simet.v9i1.1914>
- Paramitha, N. Y., Nuryaman, A., Faisol, A., & Setiawan, E. (2023). *Klasifikasi Penyakit Stroke Menggunakan Metode Naïve Bayes*. 04(01).
- Pudiastuti, R. D. (2015). *Penyakit Pemicu Stroke* (3rd ed.). Yogyakarta : Nuha Medika. <https://katalogdinarpuspurworejo.perpusnas.go.id/detail-opac?id=10057&tipe=koleksi>

- Pujiastuti, D., & Azaria, A. D. (2018). Pentingnya Siriraj Stroke Score Di Area Keperawatan Gawat Darurat. *Jurnal Kesehatan*, 5(1), 8–14.
<https://doi.org/10.35913/jk.v5i1.75>
- Putra, A. F., Yudistira, N., & Saifuddin, A. (2021). Algoritma Decision Tree Dan Smote Untuk Klasifikasi Serangan Jantung Miokarditis Yang Imbalance. *Jurnal Litbang Edusaintech*, 2(2), 112–122.
<https://doi.org/10.51402/jle.v2i2.48>
- Renuka, D. K., Visalakshi, P., & Sankar, T. (2015). Improving E-Mail Spam Classification using Ant Colony Optimization Algorithm. *International Journal of Computer Applications*, 22–16.
- Rumagit, R. Y. (2019). *Imbalanced Dataset*.
<https://socs.binus.ac.id/2019/12/26/imbalanced-dataset/>
- Sari, M. K., Ernawati, & Wisnubhadra, I. (2016). Pembangunan Aplikasi Klasifikasi Mahasiswa Baru untuk Prediksi Hasil Studi Menggunakan Naïve Bayes Classifier. *Jurnal Buana Informatika*, 7(2), 135–142.
<https://doi.org/10.24002/jbi.v7i2.492>
- Setiawan, P. A. (2021). Diagnosis Dan Tatalaksana Stroke Hemoragik. *Jurnal Medika Hutama*, 3(1), 1660–1665.
- Suarga. (2012). *Algoritma dan Pemrograman* (2nd ed.). Yogyakarta: Andi.
<https://inlislite.uin-suska.ac.id/opac/detail-opac?id=15345>
- Utomo, D. P., & Mesran, M. (2020). Analisis Komparasi Metode Klasifikasi Data Mining dan Reduksi Atribut Pada Data Set Penyakit Jantung. *JURNAL*

MEDIA INFORMATIKA BUDIDARMA, 4(2), 437.

<https://doi.org/10.30865/mib.v4i2.2080>

Watratan, A. F., B, A. Puspita., & Moeis, D. (2020). Implementasi Algoritma Naive Bayes Untuk Memprediksi Tingkat Penyebaran Covid-19 Di Indonesia.

Journal of Applied Computer Science and Technology, 1(1), 7–14.

<https://doi.org/10.52158/jacost.v1i1.9>

Yeh, D.-Y., Cheng, C.-H., & Chen, Y.-W. (2011). A predictive model for cerebrovascular disease using data mining. *Expert Syst. Appl.*, 38, 8970–

8977. <https://doi.org/10.1016/j.eswa.2011.01.114>

Yusa, H., Sudarma, M., & Pramaita, N. (2018). HR Potensi Pelanggan Tunggalan PDAM Menggunakan Metode K-Medoids dengan Optimasi Ant Colony

Optimization (ACO). *Majalah Ilmiah Teknologi Elektro*, 17(3), 353.

<https://doi.org/10.24843/MITE.2018.v17i03.P08>