

OPTIMASI ALGORITMA *NAÏVE BAYES* MENGGUNAKAN
ARTIFICIAL BEE COLONY UNTUK KLASIFIKASI DATA
PENDERITA PENYAKIT STROKE

*Diajukan Untuk Menyusun Skripsi
di Jurusan Teknik Informatika Fakultas Ilmu Komputer UNSRI*



Oleh :

Melsra Safebri

NIM : 09021282025055

Jurusan Teknik Informatika

FAKULTAS ILMU KOMPUTER UNIVERSITAS SRIWIJAYA

2024

LEMBAR PENGESAHAN SKRIPSI

OPTIMASI ALGORITMA *NAÏVE BAYES* MENGGUNAKAN *ARTIFICIAL BEE COLONY* UNTUK KLASIFIKASI DATA PENDERITA PENYAKIT STROKE

Oleh :

Melsra Safebri
NIM : 09021282025055

Indralaya, 27 Mei 2024

Mengetahui,
Ketua Jurusan Teknik Informatika



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

Pembimbing,



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

TANDA LULUS UJIAN KOMPREHENSIF SKRIPSI

Pada hari Senin tanggal 20 Mei 2024 telah dilaksanakan ujian komprehensif skripsi oleh Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya

Nama : Melsra Safebri
NIM : 09021282025055
Judul : Optimasi Algoritma *Naïve Bayes* Menggunakan *Artificial Bee Colony* Untuk Klasifikasi Data Penderita Penyakit

dan dinyatakan LULUS.

1. Ketua Penguji

Osvari Arsalan, M.T.
NIP. 198806282018031001



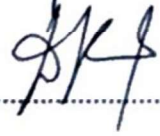
2. Penguji I

Rizki Kurniati, M.T.
NIP. 199107122019032016



3. Pembimbing I

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



Mengetahui,
Ketua Jurusan Teknik Informatika



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

HALAMAN PERNYATAAN BEBAS PLAGIASI

Yang bertanda tangan di bawah ini:

Nama : Melsra Safebri

NIM : 09021282025055

Jurusan : Teknik Informatika

Judul : Optimasi Algoritma *Naïve Bayes* Menggunakan *Artificial Bee Colony* Untuk Klasifikasi Data Penderita Penyakit Stroke

Hasil pengecekan iThenticate/Turnitin : 17%

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

Dengan pernyataan ini saya buat dengan sebenarnya dan tidak ada paksaan oleh siapa pun.



Indralaya, 27 Mei 2024



Melsra Safebri
NIM. 09021282025055

MOTTO DAN PERSEMBAHAN

“ You don’t have to be **great** to **start**, but
you have to **start** to be **great**”

— Zig Ziglar

-

Kupersembahkan Karya Tulis ini kepada:

- Diri Sendiri
- Kedua Orang Tua
- Adikku
- Sahabat dan Teman Seperjuangan
- Teknik Informatika Reguler 2020
- Para Dosen dan Dosen Pembimbing
- Fakultas Ilmu Komputer Universitas
Sriwijaya

ABSTRACT

This research aims to optimize the Naïve Bayes algorithm for feature selection of stroke disease patient data using the Artificial Bee Colony (ABC) method. Naïve Bayes is one of the simple yet effective classification algorithms. However, Naïve Bayes performance can decrease when there are irrelevant features in the data. To overcome this problem, a feature selection method with Artificial Bee Colony is applied. ABC is a population-based optimization algorithm that mimics the food search behavior of honey bees with employed bee, onlooker bee, and scout bee types, which are effective in finding optimal solutions. In this study, ABC is used to select the most relevant feature subset, thereby improving the classification accuracy of Naïve Bayes. The results show that Naïve Bayes has an accuracy of 86.10% without oversampling, while with feature selection using Artificial Bee Colony, the accuracy increases to 95.20% without oversampling. With the application of oversampling, the accuracy of Naïve Bayes reaches 82.23%, while the accuracy with Artificial Bee Colony optimization is 84.5%. The experimental results show that the combination of Naïve Bayes with ABC significantly improves the prediction accuracy compared to Naïve Bayes without feature selection. This approach is effective for stroke disease diagnosis.

Keywords : *Artificial Bee Colony, Classification, Feature Selection, Naïve Bayes, Stroke*

ABSTRAK

Penelitian ini bertujuan mengoptimasi algoritma *Naïve Bayes* untuk seleksi fitur data penderita penyakit stroke dengan menggunakan metode *Artificial Bee Colony* (ABC). *Naïve Bayes* merupakan salah satu algoritma klasifikasi yang sederhana namun efektif. Namun, performa *Naïve Bayes* dapat menurun ketika terdapat fitur-fitur yang tidak relevan dalam data. Untuk mengatasi masalah ini, metode seleksi fitur dengan *Artificial Bee Colony* diterapkan. ABC adalah algoritma optimasi berbasis populasi yang meniru perilaku pencarian makanan pada lebah madu dengan jenis *employed bee*, *onlooker bee*, dan *scout bee*, yang efektif dalam menemukan solusi optimal. Dalam penelitian ini, ABC digunakan untuk memilih subset fitur yang paling relevan, sehingga meningkatkan akurasi klasifikasi *Naïve Bayes*. Hasil penelitian menunjukkan bahwa *Naïve Bayes* memiliki akurasi 86,10% tanpa *oversampling*, sedangkan dengan seleksi fitur menggunakan *Artificial Bee Colony*, akurasi meningkat menjadi 95,20% tanpa *oversampling*. Dengan penerapan *oversampling*, akurasi *Naïve Bayes* mencapai 82,23%, sementara akurasi dengan optimalisasi *Artificial Bee Colony* adalah 84,5%. Hasil eksperimen menunjukkan bahwa kombinasi *Naïve Bayes* dengan ABC secara signifikan meningkatkan akurasi prediksi dibandingkan dengan *Naïve Bayes* tanpa seleksi fitur. Pendekatan ini efektif untuk diagnosa penyakit stroke.

Kata Kunci : *Artificial Bee Colony*, Klasifikasi, *Naïve Bayes*, Seleksi Fitur, Stroke

KATA PENGANTAR

Puji dan syukur atas kehadiran Allah Yang Maha Esa yang telah memberikan berkat, rahmat, dan petunjuk-Nya sehingga penulis dapat menyelesaikan tugas akhir ini dapat dengan baik. Penyelesaian tugas akhir guna memenuhi salah satu syarat menyelesaikan pendidikan program Strata-1 Program Studi Teknik Informatika pada Fakultas Ilmu Komputer di Universitas Sriwijaya.

Dalam penulisan tugas akhir ini, penulis menyadari bahwa tanpa adanya bimbingan, bantuan serta dukungan maupun petunjuk dari semua pihak, tidak mungkin tugas akhir ini dapat diselesaikan. Oleh karena itu pada kesempatan ini penulis menyampaikan rasa terima kasihnya kepada :

1. Tuhan Yang Maha Esa yang telah memberikan kesehatan, kecerdasan, kemudahan, dan kelancaran sehingga penulis dapat menyelesaikan tugas akhir dengan baik
2. Papa, Mama, Melania dan Kornelius sebagai keluarga penulis yang terus mendoakan dan mendukung penulis agar selalu diberi kemudahan dalam menjalankan dunia perkuliahan
3. Bapak Prof. Dr. Erwin, S.Si., M.Si. selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
4. Bapak Dr. M. Fachrurrozi, S.Si, M.T selaku Ketua Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya

5. Ibu Dian Palupi Rini, M.Kom., Ph.D., selaku Dosen Pembimbing yang memberikan waktu, masukan, serta motivasi sehingga penulis dapat menyelesaikan tugas akhir ini dengan baik
6. Ibu Rizki Kurniati, M.T. selaku Dosen Pembimbing Akademik selama masa perkuliahan
7. Seluruh Dosen Jurusan Teknik Informatika yang telah mengajar sehingga penulis mendapatkan ilmu yang bermanfaat selama masa perkuliahan
8. Seluruh staf administrasi dan pegawai Fakultas Ilmu Komputer yang telah membantu dalam mengurus hal administratif
9. Teman-teman TIREG 2020, Kurnia, Afifah, Betha, Dini, Citra, Euis, dan Taveto yang telah mengisi masa-masa penulis selama perkuliahan sehingga penulis bisa melewatinya dengan baik
10. Teman-teman baikku, Nehemia, Agatha, Jeannette, dan anak uhuy lainnya yang memberikan dukungan kepada penulis serta menjadi tempat bercerita penulis
11. BPH INTEL 2022 serta anggota lainnya yang telah mengajarkan penulis berbagai macam *skills* terutama dalam bidang keorganisasian
12. Tim Kampus Merdeka yang memberikan kesempatan kepada penulis untuk mengikuti kegiatan studi independen dan magang sehingga penulis mendapatkan banyak ilmu serta pengalaman berharga
13. Seluruh pihak yang tidak dapat disebutkan satu persatu yang telah membantu dalam kelancaran penyelesaian tugas akhir ini

Penulis menyadari bahwa tugas akhir ini masih banyak kekurangan karena keterbatasan pengetahuan dan pengalaman. Penulis berharap tugas akhir ini bisa menjadi pembelajaran dan berguna bagi pembaca.

Palembang, 07 April 2024

Melsra Safebri

DAFTAR ISI

Halaman

HALAMAN JUDUL.....	i
LEMBAR PENGESAHAN SKRIPSI.....	ii
TANDA LULUS UJIAN KOMPREHENSIF SKRIPSI	iii
HALAMAN PERNYATAAN BEBAS PLAGIASI	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 Masalah.....	I-1
1.3 Rumusan Masalah	I-4
1.4 Tujuan Penelitian.....	I-4
1.5 Manfaat Penelitian.....	I-5
1.6 Batasan Masalah.....	I-5
1.7 Sistematika Penulisan.....	I-6
1.8 Kesimpulan.....	I-7
BAB II KAJIAN LITERATUR.....	II-1
2.1 Pendahuluan	II-1
2.2 Landasan Teori	II-1
2.2.1 Penyakit Stroke	II-1

2.2.2	<i>Naïve Bayes</i>	II-2
2.2.3	Algoritma <i>Artificial Bee Colony</i>	II-3
2.2.4	<i>Naïve Bayes Artificial Bee Colony</i>	II-7
2.2.5	Klasifikasi	II-8
2.2.6	Seleksi Fitur	II-8
2.3	Penelitian Lain Yang Relevan.....	II-9
2.4	Kesimpulan.....	II-9
BAB III METODOLOGI PENELITIAN.....		III-1
3.1	Pendahuluan	III-1
3.2	Pengumpulan Data.....	III-1
3.3	Tahapan Penelitian.....	III-3
3.3.1	Menetapkan Kerangka Kerja.....	III-4
3.3.2	Menetapkan Kriteria Pengujian.....	III-5
3.3.3	Menetapkan Format Data Pengujian.....	III-5
3.3.4	Menentukan Alat Bantu Penelitian.....	III-6
3.3.5	Melakukan Pengujian Penelitian.....	III-7
3.3.6	Mengevaluasi Hasil Penelitian dan Membuat Kesimpulan	III-7
3.4	Metode Pengembangan Perangkat Lunak	III-8
3.4.1	Fase Insepsi	III-8
3.4.2	Fase Elaborasi	III-9
3.4.3	Fase Konstruksi	III-10
3.4.4	Fase Transisi.....	III-10
3.5	Kesimpulan.....	III-11
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	Analisis Kebutuhan	IV-2
4.2.1.3	Analisis dan Desain.....	IV-3
4.2.1.3.1	Analisis Kebutuhan Perangkat Lunak	IV-3

4.2.1.3.2	Analisis Data	IV-3
4.2.1.3.3	Analisis <i>Pre-processing</i> Data Training	IV-4
4.2.1.3.4	Analisis Klasifikasi dengan Metode <i>Naïve Bayes</i>	IV-8
4.2.1.3.5	Analisis Klasifikasi dengan Metode NBABC	IV-5
4.2.1.3.6	Perancangan Perangkat Lunak	IV-13
4.2.2	Fase Elaborasi	IV-15
4.2.2.1	Pemodelan Bisnis	IV-16
4.2.2.1.1	Perancangan Data	IV-16
4.2.2.1.2	Perancangan Antarmuka.....	IV-16
4.2.2.2	Kebutuhan Sistem	IV-17
4.2.2.3	Diagram.....	IV-18
4.2.2.3.1	Diagram Aktivitas	IV-18
4.2.2.3.2	Diagram <i>Sequence</i>	IV-19
4.2.3	Fase Konstruksi.....	IV-21
4.2.3.1	Kebutuhan Sistem	IV-21
4.2.3.2	Diagram Kelas.....	IV-21
4.2.3.3	Implementasi.....	IV-22
4.2.3.3.1	Implementasi Kelas	IV-22
4.2.3.3.2	Implementasi Antarmuka	IV-22
4.2.4	Fase Transisi.....	IV-23
4.2.4.1	Pemodelan Bisnis	IV-23
4.2.4.2	Kebutuhan Sistem	IV-24
4.2.4.3	Rencana Pengujian.....	IV-24
4.2.4.3.1	Rencana Pengujian <i>Use Case</i> Melakukan Input Data	IV-24
4.2.4.3.2	Rencana Pengujian <i>Use Case</i> Melakukan Klasifikasi <i>Naïve Bayes</i> dan <i>hybrid</i> seleksi fitur NBABC.....	IV-24
4.2.4.3.3	Implementasi	IV-25
4.2.4.3.4	Pengujian <i>Use Case</i> Melakukan Input Data.....	IV-25
4.2.4.3.5	Pengujian <i>Use Case</i> Melakukan Klasifikasi <i>Naïve Bayes</i> dan <i>hybrid</i> seleksi fitur NBABC	IV-25
4.3	Kesimpulan.....	IV-26

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 Menggunakan <i>Naïve Bayes</i>	V-1
5.2.3 Data Hasil Konfigurasi Percobaan <i>Naïve Bayes</i> dan NBABC.....	V-4
5.3 Analisis Hasil Penelitian.....	V-9
5.4 Kesimpulan.....	V-9
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	xxii

DAFTAR TABEL

Halaman

Tabel III-1. Dataset Stroke	III-2
Tabel III-2. Atribut Dataset Stroke	III-3
Tabel III-3. Rancangan Tabel Konfigurasi Pengujian Naïve Bayes	III-5
Tabel III-4. Rancangan Tabel Konfigurasi Pengujian NBABC	III-6
Tabel III-5. Perangkat Keras yang Digunakan.....	III-6
Tabel III-6. Perangkat Lunak yang Digunakan.....	III-6
Tabel III-7. Rancangan Tabel Konfigurasi Pengujian Berdasarkan Nilai Akurasi	III-8
Tabel III-8. Rancangan Tabel Perbandingan Hasil Akhir Pengujian Klasifikasi Naïve Bayes dengan NBABC	III-8
Tabel IV-1. Kebutuhan Fungsional.....	IV-2
Tabel IV-2. Kebutuhan Non-Fungsional	IV-2
Tabel IV-3. Data Training Penderita Penyakit Stroke	IV-4
Tabel IV-4. Data Testing Penderita Penyakit Stroke	IV-5
Tabel IV-5. Normalisasi pada atribut.....	IV-6
Tabel IV-6. Hasil Normalisasi pada Data Training	IV-7
Tabel IV-7. Hasil Normalisasi Pada Data Testing	IV-7
Tabel IV-8. Frekuensi Nilai pada Tiap Atribut	IV-8
Tabel IV-9. Perhitungan Probabilitas pada Tiap Atribut	IV-8
Tabel IV-10. Perhitungan Testing untuk Nilai Stroke	IV-9
Tabel IV-11. Perhitungan Testing pada Nilai Tidak Stroke	IV-9
Tabel IV-12. Perhitungan Akurasi Klasifikasi Naïve Bayes	IV-5
Tabel IV-13. Inisialisasi Populasi Secara Acak	IV-5
Tabel IV-14. Hasil Perhitungan Inisialisasi Populasi 1	IV-6
Tabel IV-15. Hasil Perhitungan Inisialisasi Populasi 2	IV-6
Tabel IV-16. Hasil Perhitungan Inisialisasi Populasi 3	IV-7
Tabel IV-17. Perhitungan Akurasi Fase Inisialisasi untuk Populasi 1	IV-7
Tabel IV-18. Perhitungan Akurasi Fase Inisialisasi untuk Populasi 2.....	IV-7
Tabel IV-19. Perhitungan Akurasi Fase Inisialisasi untuk Populasi 3.....	IV-8
Tabel IV-20. Populasi Fase Employed bee.....	IV-8
Tabel IV-21. Perhitungan Akurasi Fase Employed bee Populasi 1	IV-8
Tabel IV-22. Perhitungan Akurasi Fase Employed bee Populasi 2.....	IV-9
Tabel IV-23. Perhitungan Akurasi Fase Employed bee Populasi 3.....	IV-9
Tabel IV-24. Perbandingan Hasil Fase Inisialisasi dan Employed bee	IV-9
Tabel IV-25. Hasil Perhitungan Probabilitas	IV-10
Tabel IV-26. Hasil Roulette Wheel	IV-10
Tabel IV-27. Populasi Fase Onlooker bee	IV-10
Tabel IV-28. Perhitungan Akurasi Fase Onlooker bee Populasi 1	IV-10

Tabel IV-29. Perhitungan Akurasi Fase Onlooker bee Populasi 2	IV-11
Tabel IV-30. Perhitungan Akurasi Fase Onlooker bee Populasi 3	IV-11
Tabel IV-31. Perbandingan Hasil Fase Employed bee dan Onlooker bee.....	IV-11
Tabel IV-32. Hasil Populasi Perbandingan Sebelumnya dengan Fase Onlooker bee	IV-11
Tabel IV-33. Hasil Fase Scout bee	IV-12
Tabel IV-34. Solusi terbaik (global best).....	IV-12
Tabel IV-35. Definisi Pengguna	IV-13
Tabel IV-36. Definisi Diagram use Case.....	IV-14
Tabel IV-37. Skenario Melakukan Input Data.....	IV-14
Tabel IV-38. Skenario Melakukan Klasifikasi Naïve Bayes dan hybrid seleksi fitur NBABC	IV-15
Tabel IV-39. Implementasi Kelas	IV-22
Tabel IV-40. Rencana Pengujian Use Case Melakukan Input Data	IV-24
Tabel IV-41. Rencana Pengujian Use Case Melakukan Klasifikasi Naïve Bayes dan hybrid seleksi fitur NBABC	IV-24
Tabel IV-42. Pengujian Use Case Melakukan Input Data.....	IV-25
Tabel IV-43. Pengujian Use Case Melakukan Klasifikasi Naïve Bayes dan hybrid seleksi fitur NBABC	IV-25
Tabel V-1. Hasil Pengujian Rasio Split Data pada Klasifikasi Naïve Bayes.....	V-2
Tabel V-2. Hasil Evaluasi Naïve Bayes Sebelum Resampling.....	V-2
Tabel V-3. Hasil Evaluasi Naïve Bayes Setelah Resampling	V-3
Tabel V-4. Hasil Evaluasi NBABC dengan Jumlah Lebah = 10	V-4
Tabel V-5. Hasil Evaluasi NBABC dengan Jumlah Lebah = 20	V-5
Tabel V-6. Hasil Evaluasi NBABC dengan Jumlah Lebah = 30	V-6
Tabel V-7. Hasil Evaluasi NBABC dengan Jumlah Lebah = 40	V-7
Tabel V-8. Hasil Evaluasi NBABC dengan Jumlah Iterasi = 50	V-8
Tabel V-9. Hasil Pengujian Perbandingan Nilai Akurasi.....	V-9
Tabel V-10. Hasil Akhir Akurasi Naïve Bayes dan NBABC.....	V-9

DAFTAR GAMBAR

Halaman

Gambar III-1. Diagram Tahapan Penelitian	III-4
Gambar III-2. Tahapan Pengujian Penelitian.....	III-7
Gambar IV-1. Diagram Use Case	IV-13
Gambar IV-2. Rancangan Antarmuka Perangkat Lunak	IV-17
Gambar IV-3. Diagram Aktivitas Melakukan Input Data	IV-18
Gambar IV-4. Diagram Aktivitas Melakukan Klasifikasi Naïve Bayes dan hybrid seleksi fitur NBABC	IV-19
Gambar IV-5. Diagram Sequence Melakukan Input Data.....	IV-20
Gambar IV-6. Diagram Sequence Melakukan Klasifikasi Naïve Bayes dan hybrid seleksi fitur NBABC	IV-20
Gambar IV-7. Diagram Kelas.....	IV-21
Gambar IV-8 Antarmuka Klasifikasi Naïve Bayes + ABC	IV-23

DAFTAR LAMPIRAN

	Halaman
Lampiran 1. Dataset	xxii
Lampiran 2. Kode Program.....	xxiii

BAB I

PENDAHULUAN

1.1 Pendahuluan

Pemahaman mengenai penggunaan algoritma *Artificial Bee Colony* dalam mengoptimalkan algoritma *Naïve Bayes* untuk pengelompokan data dibahas pada bab ini. Ini mencakup latar belakang dari penggabungan dua metode tersebut, rumusan dari masalah, tujuan dari penelitian, batasan-batasan, manfaat yang diharapkan dari penelitian ini, serta terdapat kesimpulan di akhir.

1.2 Latar Belakang Masalah

Secara global, stroke adalah penyebab kematian dan kecacatan paling umum. Berdasarkan data 2019, WHO menyatakan bahwa stroke merupakan kematian tertinggi kedua setelah iskemik jantung atau dikenal dengan penyakit jantung. Stroke yang menduduki nomor dua sebagai penyebab kematian tertinggi, menyumbang sebesar 11% dari keseluruhan kematian.¹ Secara global, jumlah kasus dan kematian akibat stroke meningkat dari 4,07 juta pada jumlah kasus dan 2,05 juta pada kematian di tahun 1990 menjadi 7,86 juta pada jumlah kasus dan 3,15 juta pada kematian di tahun 2020. Jumlah kasus stroke diperkirakan meningkat menjadi 9,62 juta pada tahun 2030 (Pu et al., 2023).

Diperlukan perhatian yang serius terhadap isu stroke di Indonesia karena jumlah kasus yang terus mengalami peningkatan. Menurut Kemenkes RI pada

¹ WHO, "The top 10 causes of death" <https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death> (diakses pada tanggal 7 Oktober 2023)

Riset Kesehatan Dasar (Riskesdas) 2018, angka kejadian stroke meningkat menjadi 10,9% dari Riskesdas 2013 sebesar 7% untuk umur > 15 tahun. (Kemenkes RI, 2018)².

Stroke dapat menyebabkan kematian serta menjadi ancaman serius jika tidak ditangani dengan cepat. Hal ini terjadi karena adanya gangguan aliran peredaran darah ke otak, sehingga jika tidak ditangani dengan cepat maka kekurangan oksigen ke otak dan mengakibatkan kematian sel saraf (dr. Rizaldy Pinzon & SpS, 2010). Proses penanganan stroke memerlukan waktu yang cukup lama dan berakibat fatal jika terjadi keterlambatan dalam mendiagnosis. Untuk kasus ini diperlukan sistem yang dapat mempercepat atau menyederhanakan proses pendeteksian untuk mengurangi jumlah penderita penyakit stroke (Hutama et al., 2018). Klasifikasi menjadi salah satu cara memprediksi penyakit stroke. Hasil dari klasifikasi yang akurat akan membantu para tenaga medis dalam penanganan pasien stroke.

Salah satu metode klasifikasi yang paling umum digunakan adalah *Naïve Bayes*. Untuk metode klasifikasi ini, probabilitas dihitung dengan menggabungkan nilai-nilai dan frekuensi dataset sehingga terbilang pendekatan probabilitas sederhana (Doni et al., 2021). Salah satu kelebihan dalam menggunakan *Naïve Bayes* yaitu metode ini memerlukan jumlah data pelatihan yang kecil dalam proses klasifikasi (Saleh, 2015).

² Kemenkes RI “Hasil Utama Riskesdas 2018” https://kesmas.kemkes.go.id/assets/upload/dir_519d41d8cd98f00/files/Hasil-riskesdas-2018_1274.pdf (diakses pada tanggal 29 Oktober 2023)

Dataset yang dipergunakan dalam penelitian biasanya memiliki jumlah data serta jumlah fitur yang banyak. Hal ini dapat memakan waktu yang lama serta biaya yang lebih sehingga diperlukan pemilihan fitur. Penelitian (Suhaila et al., 2022) terlihat bahwa penggunaan metode SVM yang dioptimasi menggunakan *Artificial Bee Colony* (ABC) memberikan peningkatan dalam akurasi klasifikasi data Teh. Dengan hanya menggunakan metode SVM, akurasi yang diperoleh adalah 80%, sedangkan dengan SVM-ABC, akurasi meningkat menjadi 96,7%. Hal ini menunjukkan bahwa penggunaan optimasi ABC dengan klasifikasi SVM dapat meningkatkan kinerja model.

Penelitian terkait lainnya menggunakan metode *Naïve Bayes* dengan optimasi *Adaboost* untuk menentukan penyakit stroke (Byna & Basit, 2020) mendapatkan akurasi tertinggi sebesar 98%. Hal ini menunjukkan ada peningkatan akurasi dari penggunaan hanya *Naïve Bayes* dengan tambahan optimasi *Adaboost*.

Penelitian lain yang dilakukan (Rizka, 2017) dalam mengklasifikasikan data kanker serviks menggunakan *Naïve Bayes* dengan optimasi *Artificial Bee Colony* didapatkan akurasi sebesar 93,33%. Hal ini membuktikan bahwa NBABC lebih baik dari metode *Naïve Bayes* standar yang akurasinya hanya 60%

Pada penelitian ini, *Artificial Bee Colony* digunakan dalam meningkatkan efisiensi pada proses seleksi fitur dalam *Naïve Bayes*. Salah satu algoritma yang telah digunakan untuk menemukan solusi terbaik pada masalah optimisasi numerik adalah *Artificial Bee Colony* (Pradnyana et al., 2018). Diharapkan bahwa

efektivitas proses klasifikasi data dapat ditingkatkan dengan memanfaatkan algoritma *Artificial Bee Colony* untuk metode optimasi *Naïve Bayes*.

1.3 Rumusan Masalah

Rumusan masalah penelitian ini berdasarkan penjelasan latar belakang yang telah diuraikan sebelumnya adalah sebagai berikut:

1. Bagaimana merancang perangkat lunak untuk klasifikasi data penderita penyakit stroke menggunakan metode *Naïve Bayes* yang dioptimalkan dengan algoritma *Artificial Bee Colony*?
2. Bagaimana hasil klasifikasi data penderita penyakit stroke dengan menerapkan metode *Naïve Bayes* yang telah dioptimalkan menggunakan algoritma *Artificial Bee Colony*?

1.4 Tujuan Penelitian

Berikut ini adalah tujuan penelitian:

1. Merancang perangkat lunak yang mampu melakukan klasifikasi data penderita penyakit stroke menggunakan metode *Naïve Bayes* yang telah dioptimalkan dengan algoritma *Artificial Bee Colony*. Tujuan ini mencakup pengembangan sistem yang efisien dan akurat dalam mengidentifikasi kemungkinan kasus penyakit stroke berdasarkan data yang tersedia.
2. Mengevaluasi hasil klasifikasi data penderita penyakit stroke dengan menerapkan metode *Naïve Bayes* yang telah dioptimalkan

menggunakan algoritma *Artificial Bee Colony*. Tujuan ini mencakup analisis performa sistem klasifikasi yang dikembangkan, termasuk akurasi dan metrik evaluasi lainnya untuk menilai seberapa baik sistem dapat mengklasifikasikan kasus-kasus penyakit stroke.

1.5 Manfaat Penelitian

Berikut ini adalah manfaat penelitian:

1. Meningkatkan akurasi dalam proses klasifikasi data, sehingga memungkinkan pengambilan keputusan yang lebih akurat dalam diagnosis atau prediksi penyakit stroke.
2. Memahami bagaimana algoritma *Naïve Bayes* menerapkan optimasi menggunakan *Artificial Bee Colony* untuk melakukan klasifikasi data penderita penyakit stroke.

1.6 Batasan Masalah

Berikut ini adalah batasan masalah yang diterapkan dalam penelitian:

1. Menggunakan data sekunder yaitu dataset penderita penyakit stroke yang diunduh dari situs *kaggle*³
2. Berkas yang akan dimasukkan harus dalam format .csv.

³ “Stroke Prediction Dataset”, <https://www.kaggle.com/datasets/fedesoriano/stroke-prediction-dataset/data>

1.7 Sistematika Penulisan

Tugas akhir ini disesuaikan dengan sistematika penulisan standar yang ditetapkan oleh Fakultas Ilmu Komputer Universitas Sriwijaya, yaitu:

BAB I. PENDAHULUAN

Latar belakang masalah, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, dan sistematika penulisan dibahas dalam bab ini.

BAB II. KAJIAN LITERATUR

Teori dasar yang relevan dengan penelitian akan dibahas dalam bab ini, bersama dengan temuan penelitian sebelumnya yang terkait.

BAB III. METODOLOGI PENELITIAN

Tahapan yang akan dilakukan dalam penelitian, serta suatu kerangka kerja dan perancangan manajemen proyek untuk pelaksanaannya dibahas dalam bab ini.

BAB IV. PENGEMBANGAN PERANGKAT LUNAK

Metode *Rational Unified Process* (RUP) dibahas dalam bab ini, digunakan untuk perancangan dan pengembangan perangkat lunak.

BAB V. HASIL DAN ANALISIS PENELITIAN

Bab ini membahas mengenai hasil pengujian dan penelitian sesuai dengan tahapan yang telah direncanakan sebelumnya. Kesimpulan penelitian akan didasarkan pada analisis hasil dari penelitian ini.

BAB VI. KESIMPULAN DAN SARAN

Kesimpulan dari hasil penelitian dan saran untuk pengembangan penelitian selanjutnya dibahas pada bab ini.

1.8 Kesimpulan

Masalah pada penelitian terkait peningkatan hasil kinerja algoritma *Naïve Bayes* dibahas pada bab ini. Algoritma *Artificial Bee Colony* diimplementasikan ke dalam *Naïve Bayes* untuk data penyakit stroke.

DAFTAR PUSTAKA

- Akay, B., & Karaboga, D. (2012). A modified *Artificial Bee Colony* algorithm for real-parameter optimization. *Information Sciences*, 192, 120–142. <https://doi.org/10.1016/j.ins.2010.07.015>
- Bolaji, A. L., Khader, A. T., Al-Betar, M. A., & Awadallah, M. A. (2005). *Artificial Bee Colony Algorithm, Its Variants And Applications: A Survey*. . . Vol., 47.
- Borgelt, C. (1999). *A Naïve Bayes Classifier Plug-In for DataEnginetm*.
- Byna, A., & Basit, M. (2020). Penerapan Metode Adaboost Untuk Mengoptimasi Prediksi Penyakit Stroke Dengan Algoritma Naïve Bayes. *Jurnal Sisfokom (Sistem Informasi Dan Komputer)*, 9(3), 407–411. <https://doi.org/10.32736/sisfokom.v9i3.1023>
- Dinata, C. A., Safrita, Y. S., & Sastri, S. (2013). Gambaran Faktor Risiko dan Tipe Stroke pada Pasien Rawat Inap di Bagian Penyakit Dalam RSUD Kabupaten Solok Selatan Periode 1 Januari 2010—31 Juni 2012. *Jurnal Kesehatan Andalas*, 2(2), 57. <https://doi.org/10.25077/jka.v2i2.119>
- Dini, D. R. W. (2022). *Penerapan Metode Backpropagation Dan Algoritma Bee Colony Pada Prediksi Penyakit Diabetes Melitus*. Universitas Sriwijaya.
- Doni, B. T. R., Susanti, S., & Mubarak, A. (2021). Penerapan Data Mining Untuk Klasifikasi Penyakit Hepatocellular Carcinoma Menggunakan Algoritma Naïve Bayes. *Jurnal Responsif: Riset Sains Dan Informatika*, 3(1), Article 1. <https://doi.org/10.51977/jti.v3i1.403>
- dr. Rizaldy Pinzon, M. S. S., & SpS, L. A. (2010). *Awas Stroke! Pengertian, Gejala, Tindakan, Perawatan dan Pencegahan*. Penerbit Andi. <https://books.google.co.id/books?id=TrFtdwJ8qwkC>
- Gao, W., & Liu, S. (2012). A modified *Artificial Bee Colony* algorithm. *Computers & Operations Research*, 39(3), 687–697. <https://doi.org/10.1016/j.cor.2011.06.007>

- Guo, H., Zhang, L., Ren, Y., Meng, L., Zhou, Z., & Li, J. (2022). Neighborhood Modularization-based *Artificial Bee Colony* Algorithm for Disassembly Planning with Operation Attributes. *Chinese Journal of Mechanical Engineering*, 35(1), 143. <https://doi.org/10.1186/s10033-022-00812-2>
- 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), Article 11.
- Kalouisis, A., Prados, J., & Hilario, M. (2007). Stability of feature selection algorithms: A study on high-dimensional spaces. *Knowledge and Information Systems*, 12(1), 95–116. <https://doi.org/10.1007/s10115-006-0040-8>
- Karaboga, D., & Basturk, B. (2007). A powerful and efficient algorithm for numerical function optimization: *Artificial Bee Colony* (ABC) algorithm. *Journal of Global Optimization*, 39(3), 459–471. <https://doi.org/10.1007/s10898-007-9149-x>
- Kemenkes RI. (2018). *Laporan Nasional Riskesdas 2018*. Badan Penelitian dan Pengembangan Kesehatan.
- Kesavaraj, G., & Sukumaran, S. (2013). A study on classification techniques in data mining. *2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT)*, 1–7. <https://doi.org/10.1109/ICCCNT.2013.6726842>
- Kumar, V. (2014). Feature Selection: A literature Review. *The Smart Computing Review*, 4(3). <https://doi.org/10.6029/smartcr.2014.03.007>
- Neelamegam, S., & Ramaraj, D. E. (2013). *Classification algorithm in Data mining: An Overview*. 4(8).
- Nurdina, A., & Puspita, A. B. I. (2023). *Naïve Bayes* and KNN for Airline Passenger Satisfaction Classification: Comparative Analysis. *Journal of Information System Exploration and Research*, 1(2). <https://doi.org/10.52465/joiser.v1i2.167>

- Pradnyana, I. P. B. A., Soebroto, A. A., & Perdana, R. S. (2018). Peramalan Curah Hujan Menggunakan Metode Jaringan Saraf Tiruan Dengan Optimasi Algoritma Bee Colony. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 2(10), Article 10.
- Pu, L., Wang, L., Zhang, R., Zhao, T., Jiang, Y., & Han, L. (2023). Projected Global Trends in Ischemic Stroke Incidence, Deaths and Disability-Adjusted Life Years From 2020 to 2030. *Stroke*, 54(5), 1330–1339. <https://doi.org/10.1161/STROKEAHA.122.040073>
- Rizka, D. I. (2017). *Klasifikasi Data Kanker Serviks Menggunakan Metode Naïve Bayes dengan Pemilihan Fitur Artificial Bee Colony*. Universitas Indonesia.
- Saleh, A. (2015). *Implementasi Metode Klasifikasi Naïve Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga*. 2(3).
- Suhaila, S., Lelono, D., & Sari, Y. S. (2022). Seleksi Fitur dengan *Artificial Bee Colony* untuk Optimasi Klasifikasi Data Teh menggunakan Support Vector Machine. *IJEIS (Indonesian Journal of Electronics and Instrumentation Systems)*, 12(1), 81. <https://doi.org/10.22146/ijeis.63902>
- TSai, P.-W., Pan, J.-S., Liao, B.-Y., & Chu, S.-C. (2009). Enhanced *Artificial Bee Colony* Optimization. *International Journal of Innovative Computing, Information and Control*, 5(12), 5081–5092.
- Tu, W.-J., Wang, L.-D., on behalf of the Special Writing Group of China Stroke Surveillance Report, Yan, F., Peng, B., Hua, Y., Liu, M., Ji, X.-M., Ma, L., Shan, C.-L., Wang, Y.-L., Zeng, J.-S., Chen, H.-S., Fan, D.-S., Gu, Y.-X., Tan, G.-J., Hu, B., Kang, D.-Z., Liu, J.-M., ... Wu, J. (2023). China stroke surveillance report 2021. *Military Medical Research*, 10(1), 33. <https://doi.org/10.1186/s40779-023-00463-x>
- Yilmaz, A. B., Taspinar, Y. S., & Koklu, M. (2022). Classification of Malicious Android Applications Using *Naïve Bayes* and Support Vector Machine Algorithms. *International Journal of Intelligent Systems and Applications in Engineering*, 10(2), Article 2.