

**OPTIMASI ALGORITMA *DECISION TREE* MENGGUNAKAN
METODE *ARTIFICIAL BEE COLONY* UNTUK KLASIFIKASI
DATA PENDERITA PENYAKIT DIABETES**

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



Oleh:

Dini Andriani

NIM: 09021182025006

**Jurusan Teknik Informatika
FAKULTAS ILMU KOMPUTER UNIVERSITAS SRIWIJAYA
2024**

LEMBAR PENGESAHAN SKRIPSI

**OPTIMASI ALGORITMA *DECISION TREE* MENGGUNAKAN METODE
ARTIFICIAL BEE COLONY UNTUK KLASIFIKASI DATA PENDERITA
PENYAKIT DIABETES**

Oleh:

**Dini Andriani
NIM: 09021182025006**

Palembang, 1 Juli 2024

Mengetahui,
Ketua Jurusan Teknik Informatika



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

Pembimbing

A handwritten signature in black ink, appearing to be "DPR", is written above the name of the supervisor.

Dian Palupi Rini, M.Kom., Ph.D.
NIP 192802232006042002

TANDA LULUS UJIAN KOMPREHENSIF SKRIPSI

Pada hari Jumat tanggal 21 Juni 2024 telah dilaksanakan ujian komprehensif skripsi oleh Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya

Nama : Dini Andriani
NIM : 09021182025006
Judul : Optimasi Algoritma *Decision Tree* Menggunakan Metode *Artificial Bee Colony* Untuk Klasifikasi Data Penderita Penyakit Diabetes

dan dinyatakan

1. Ketua Penguji

Hadipurnawan Satria, Ph.D.
NIP 198004182020121001



.....

2. Penguji

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



.....

3. Pembimbing

Dian Palupi Rini., M.Kom., Ph.D.
NIP 192802232006042002



.....

Mengetahui,
Ketua Jurusan Teknik Informatika



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

HALAMAN PERNYATAAN

Yang bertanda tangan di bawah ini:

Nama : Dini Andriani

NIM : 09021182025006

Program Studi : Teknik Informatika

Judul : Optimasi Algoritma *Decision Tree* Menggunakan Metode *Artificial Bee Colony* Untuk Klasifikasi Data Penderita Penyakit Diabetes

Hasil pengecekan iThenticate/Turnitin: 13 %

Menyatakan bahwa Laporan Proyek 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, 31 Mei 2024

Dini Andriani

09021182025006

MOTTO DAN PERSEMBAHAN

مَنْ جَدَّ وَجَدَ

مَنْ صَبَرَ ظَفِرَ

مَنْ سَارَ عَلَى الدَّرْبِ وَصَلَ

...

Siapa yang bersungguh-sungguh akan berhasil

Siapa yang bersabar akan beruntung

Siapa yang berjalan di jalan-Nya akan sampai di tujuan

~Imam Syafi'i~

Kupersembahkan karya tulis ini kepada:

- Allah SWT
- Alm. Kedua orang tuaku
- Kakak-kakakku
- Keluarga besarku
- Sahabat dan teman seperjuangan
- Dosen pembimbing
- Almamater

**OPTIMIZATION OF *DECISION TREE* ALGORITHM USING
ARTIFICIAL BEE COLONY FOR DATA CLASSIFICATION OF
DIABETES DISEASE PATIENTS DATA**

By:

Dini Andriani
09021182025006

ABSTRACT

Diabetes is a long-term or chronic disease. Diabetes can attack anyone and under any conditions and will continue for a lifetime. In reality, most diabetes is predicted too late, causing complications with other diseases which will ultimately lead to death. This system is expected to be able to alert the risk of complications in diabetes patients in the future. Extract knowledge from diabetes data with machine learning to learn patterns. This research tests the effect of Decision Tree C4.5 optimization for classification with Artificial Bee Colony for selecting data attributes to be used. Classification using the Decision Tree C4.5 algorithm produces an accuracy of 0.74. Meanwhile, after selecting the Artificial Bee Colony feature, it produced an accuracy value of 0.77. The increase in classification accuracy reached 0.03. Optimization using the Artificial Bee Colony method succeeded in increasing the accuracy of the Decision Tree C4.5 algorithm in classification diabetes data.

Keywords: Classification, Decision Tree C4.5, Artificial Bee Colony, Feature Selection, Diabetes

Approved,
Head of Informatics Department



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

Palembang, 3 July 2024

Supervisor,



Dian Puspita Nini, M.Kom., Ph.D.
NIP 192802232006042002

OPTIMASI ALGORITMA *DECISION TREE* MENGGUNAKAN METODE *ARTIFICIAL BEE COLONY* UNTUK KLASIFIKASI DATA PENDERITA PENYAKIT DIABETES

Oleh:

Dini Andriani
09021182025006

ABSTRAK

Diabetes adalah penyakit yang berlangsung jangka panjang atau penyakit kronis. Penyakit diabetes dapat menyerang siapa saja dan dalam kondisi apa saja serta akan terus berlanjut atau seumur hidup. Pada kenyataannya kebanyakan penyakit diabetes terlambat diprediksi sehingga menyebabkan terjadinya risiko komplikasi pada pasien diabetes di waktu mendatang sehingga menyebabkan terjadinya komplikasi dengan penyakit lain yang pada akhirnya akan berujung pada kematian. Sistem ini diharapkan dapat mendeteksi. Mengambil pengetahuan dari data penderita diabetes dengan pembelajaran mesin untuk mempelajari polanya. Penelitian ini menguji pengaruh optimasi *Decision Tree C4.5* untuk klasifikasi dengan *Artificial Bee Colony* untuk pemilihan atribut data yang akan digunakan. Klasifikasi menggunakan algoritma *Decision Tree C4.5* menghasilkan akurasi sebesar 0.74. Sedangkan setelah dilakukan seleksi fitur *Artificial Bee Colony* menghasilkan nilai akurasi 0.77. Peningkatan akurasi klasifikasi mencapai 0.03. Optimasi menggunakan metode *Artificial Bee Colony* berhasil meningkatkan akurasi algoritma *Decision Tree C4.5* dalam melakukan klasifikasi data penyakit diabetes.

Kata Kunci: Klasifikasi, *Decision Tree C4.5*, *Artificial Bee Colony*, Seleksi Fitur, Penyakit Diabetes


Mengetahui,
Ketua Jurusan Teknik Informatika



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

Palembang, 3 Juli 2024

Pembimbing



Dian Puspita Hini., M/Kom., Ph.D.
NIP. 192802232006042002

KATA PENGANTAR

Puji dan syukur atas berkat dan rahmat Allah SWT yang telah memberikan hidayah, rahmat, dan petunjuk-Nya sehingga penulis dapat menyelesaikan tugas akhir ini, disusun guna memenuhi salah satu syarat penyelesaian pendidikan Strata-1 di Fakultas Ilmu Komputer Program Studi Teknik Informatika, Universitas Sriwijaya. Dalam penelitian dan penulisan skripsi ini, penulis tidak terlepas dari berbagai hambatan dan kesulitan yang pada dasarnya memberikan hikmah tersendiri bagi penulis.

Penulis menyadari bahwa penulisan ini tidak dapat terselesaikan tanpa dukungan dari berbagai pihak baik moril maupun material. Oleh karena itu, penulis ingin menyampaikan ucapan terima kasih kepada semua pihak yang telah membantu dalam penyusunan skripsi ini terutama kepada:

1. Allah Subhanahu wa Ta'ala yang telah memberikan, kesehatan, kecerdasan, kelancaran sehingga dapat menyelesaikan Skripsi ini.
2. Nabi Muhammad Shallallahu 'Alaihi wa sallam yang telah memberikan jalan yang terang dan ilmu yang bermanfaat pada umatnya.
3. Kedua orang tuaku, almarhum Bapak Nurhopit dan almarhumah Mamah Ikah, semoga mereka bangga dengan perjuangan anak bungsunya ini.
4. A Yadi dan A Fahmi selaku kakak penulis, serta keluarga besar yang telah memberikan yang telah memberikan dukungan dan doa untuk kelancaran selama mengerjakan skripsi ini.

5. Prof. DR. Erwin, S.Si., M.Si. selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
6. Dr. M. Fachrurrozi, S.Si., M.T. selaku Ketua Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya, Dosen Pembimbing Akademik, sekaligus Dosen Penguji tugas akhir yang telah memberikan saran dan masukan yang membangun.
7. Dian Palupi Rini, M.Kom., Ph.D. selaku Dosen Pembimbing yang telah menyediakan waktu, senantiasa memberikan bimbingan, masukan, perhatian, dan motivasi untuk menyelesaikan tugas akhir ini.
8. Seluruh Dosen Jurusan Teknik Informatika dan Dosen Fakultas Ilmu Komputer yang telah banyak memberikan ilmu selama masa perkuliahan.
9. Mba Pita, Mba Septy, Kak Ricy, Mba Dwi, Mba Tari, Kak Azhar, Kak Welli, dan seluruh staff yang telah membantu dalam kelancaran proses administrasi dan akademik selama masa perkuliahan.
10. Seseorang yang telah banyak membantu proses penelitian ini, memberi semangat untuk berusaha, dan selalu siap sedia jika penulis membutuhkan bantuan, semoga menjadi *Game Developer* terbaik.
11. KPU FASILKOM, LDF WIFI, FASCO, GDSC UNSRI, Elmu.id, Geulis Indonesia, Hactiv8, MyEduSolve, Lab Mulmed, dan Lab AI yang telah memberikan ruang bagi Penulis untuk berprestasi dan berkarya selama masa perkuliahan.
12. Kakak-kakak dan adik-adik tingkatku, teman-teman seperjuanganku khususnya kelas TI REG C, *Informatic Sixteent Generation*, anggota grup

Skripsi Garis Keras, Senja, DC, BKN, yang telah berbagi keluh kesah, semangat, tawa, cerita dan bantuan selama perkuliahan.

13. Rekan terdekatku dari jurusan SK dan SI. Fakultas FK, FKM, FKG, FKIP, FMIPA, FISIP yang tidak dapat kusebutkan satu persatu namanya.
14. Teman TK-SMA yang masih komunikasi denganku walau jarak memisahkan tapi selalu ada waktu untuk saling mendengarkan keluh kesah.
15. Serta berbagai pihak yang penulis tidak bisa sebutkan satu persatu, hanya ucapan terima kasih dari lubuk hati yang paling dalam penulis haturkan dan semoga amal dan jasa yang mulia sahabat-sahabat akan dicatat sebagai amal kebaikan dan dibalas sesuai amal perbuatannya oleh Allah Swt.

Penulis 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.

Semoga semua pihak yang telah membantu dalam penulisan skripsi ini mendapatkan pahala di sisi Allah Swt. dan penulis berharap semoga skripsi ini bermanfaat, khususnya bagi para akademisi. *Aamiin*.

Palembang, 31 Mei 2024

Dini Andriani

DAFTAR ISI

	Halaman
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.....	xi
DAFTAR TABEL.....	xv
DAFTAR GAMBAR.....	xvii
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-6
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 Diabetes.....	II-1
2.2.2 Klasifikasi.....	II-3
2.2.3 <i>Decision Tree</i>	II-4
2.2.3.1 Algoritma C4.5.....	II-5
2.2.4 <i>Artificial Bee Colony</i>	II-7
2.2.5 <i>Confusion Matrix</i>	II-12

2.3 <i>Rational Unified Process (RUP)</i>	II-13
2.4 Penelitian Lain yang Relevan	II-15
2.5 Kesimpulan	II-16
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 Kerangka Kerja Penelitian	III-6
3.3.2 Kriteria Pengujian	III-7
3.3.3 Format Pengujian	III-7
3.3.4 Alat Bantu Penelitian	III-8
3.3.5 Pengujian Penelitian	III-8
3.3.6 Analisis Hasil Pengujian dan Kesimpulan Penelitian	III-8
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-14
BAB IV PENGEMBANGAN PERANGKAT LUNAK	IV-1
4.1 Pendahuluan	IV-1
4.2 Fase Insepsi	IV-1
4.2.1 Pemodelan Bisnis	IV-1
4.2.2 Kebutuhan Sistem	IV-2
4.2.3 Analisis dan Desain	IV-3
4.2.3.1 Analisis Kebutuhan Perangkat Lunak	IV-3
4.2.3.2 Analisis Data	IV-3
4.2.3.3 Analisis <i>Decision Tree</i> C4.5	IV-4
4.2.3.4 Analisis <i>Artificial Bee colony</i>	IV-4
4.2.3.5 Desain Perangkat Lunak	IV-4

4.3 Fase Elaborasi	IV-9
4.3.1 Pemodelan Bisnis	IV-9
4.3.1.1 Perancangan Antarmuka	IV-9
4.3.1.2 Perancangan Data	IV-12
4.3.2 Kebutuhan Sistem	IV-12
4.3.3 Diagram	IV-13
4.3.3.1 Diagram Aktivasi	IV-13
4.3.3.2 <i>Sequence</i> Diagram	IV-15
4.4 Fase Konstruksi	IV-16
4.4.1 Kebutuhan Sistem	IV-16
4.4.2 Diagram Kelas	IV-17
4.4.3 Implementasi	IV-17
4.4.3.1 Implementasi Kelas	IV-17
4.4.3.2 Implementasi Antarmuka	IV-18
4.5 Fase Transisi	IV-20
4.5.1 Pemodelan Bisnis	IV-20
4.5.2 Kebutuhan Sistem	IV-20
4.5.3 Rencana Pengujian	IV-21
4.5.3.1 Rencana Pengujian <i>Use Case</i> Klasifikasi <i>Decision Tree</i> C4.5 dengan <i>Artificial Bee Colony</i>	IV-21
4.5.3.2 Rencana Pengujian Memilih File	IV-22
4.5.3.3 Rencana Pengujian Input Data Baru dan Memilih Metode untuk Prediksi	IV-22
4.5.4 Implementasi	IV-23
4.5.4.1 Pengujian <i>Use Case</i> Klasifikasi <i>Decision Tree</i> C4.5 dengan <i>Artificial Bee Colony</i>	IV-23
4.5.4.2 Pengujian Memilih File	IV-24
4.5.4.3 Pengujian Input Data Baru dan Memilih Metode untuk Prediksi	IV-25
4.6 Kesimpulan	IV-27
BAB V HASIL DAN ANALISIS PENELITIAN	V-1

5.1 Pendahuluan	V-1
5.2 Data Hasil Penelitian	V-1
5.2.1 Konfigurasi Percobaan	V-1
5.2.2 Data Hasil Konfigurasi Percobaan <i>Decision Tree</i> C4.5 dan DTABC	V-3
5.3 Analisis Hasil Penelitian	V-11
5.4 Kesimpulan	V-13
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	xxiii

DAFTAR TABEL

	Halaman
Tabel III- 1 Sampel <i>Diabetes Dataset</i>	III-2
Tabel III- 2 Atribut <i>Diabetes Dataset</i>	III-2
Tabel III- 3 Rancangan Hasil Pengujian DT dan DTABC	III-7
Tabel III- 4 Rancangan Tabel Konfigurasi Pengujian Berdasarkan Nilai Akurasi	III-9
Tabel III- 5 Rancangan Tabel Perbandingan Akhir Hasil Pengujian Klasifikasi <i>Decision Tree</i> dan DTABC	III-9
Tabel III- 6 <i>Work Breakdown Structure (WBS)</i>	III-12
Tabel IV- 1 Kebutuhan Fungsional Perangkat Lunak	IV-2
Tabel IV- 2 Non-Fungsional Perangkat Lunak	IV-3
Tabel IV- 3 Definisi Pengguna	IV-5
Tabel IV- 4 Definisi <i>Use Case</i>	IV-5
Tabel IV- 5 Skenario Klasifikasi <i>Decision Tree C4.5</i> dengan <i>Artificial Bee Colony</i>	IV-6
Tabel IV- 6 Skenario Memilih File	IV-7
Tabel IV- 7 Skenario Input Data Baru dan Memilih Metode untuk Prediksi ...	IV-8
Tabel IV- 8 Implementasi Kelas	IV-18
Tabel IV- 9 Rencana Pengujian <i>Use Case</i> Klasifikasi DT dengan ABC	IV-21
Tabel IV- 10 Rencana Pengujian Memilih File	IV-22
Tabel IV- 11 Rencana Pengujian Input Data Baru dan Memilih Metode untuk Prediksi	IV-22
Tabel IV- 12 Pengujian <i>Use Case</i> Klasifikasi <i>Decision Tree C4.5</i> dengan <i>Artificial Bee Colony</i>	IV-23
Tabel IV- 13 Pengujian <i>Use Case</i> Memilih File	IV-24
Tabel IV- 14 Pengujian <i>Use Case</i> input data baru dan memilih metode untuk prediksi	IV-25
Tabel V- 1 Hasil Evaluasi Klasifikasi dengan <i>Decision Tree C4.5</i> Tanpa <i>Resampling</i>	V-3

Tabel V- 2 Hasil Evaluasi Klasifikasi dengan <i>Decision Tree</i> C4.5	V-4
Tabel V- 3 Hasil Evaluasi DTABC dengan Bee=10 dan Iterasi=10	V-4
Tabel V- 4 Hasil Evaluasi DTABC dengan Bee=10 dan Iterasi=20	V-4
Tabel V- 5 Hasil Evaluasi DTABC dengan Bee = 10 dan Iterasi = 50	V-5
Tabel V- 6 Hasil Evaluasi DTABC dengan Bee=10, Iterasi=75	V-5
Tabel V- 7 Hasil Evaluasi DTABC dengan Bee=10 dan Iterasi=100	V-5
Tabel V- 8 Hasil Evaluasi DTABC dengan Bee=20 dan Iterasi=10	V-5
Tabel V- 9 Hasil Evaluasi DTABC dengan Bee=20 dan Iterasi=20	V-6
Tabel V- 10 Hasil Evaluasi DTABC dengan Bee=20 dan Iterasi=50	V-6
Tabel V- 11 Hasil Evaluasi DTABC dengan Bee=20 dan Iterasi=50	V-6
Tabel V- 12 Hasil Evaluasi DTABC dengan Bee=20 dan Iterasi=75	V-6
Tabel V- 13 Hasil Evaluasi DTABC dengan Bee=20 dan Iterasi=100	V-7
Tabel V- 14 Hasil Evaluasi DTABC dengan Bee=30 dan Iterasi=10	V-7
Tabel V- 15 Hasil Evaluasi DTABC dengan Bee=30 dan Iterasi=20	V-7
Tabel V- 16 Hasil Evaluasi DTABC dengan Bee=30 dan Iterasi=50	V-7
Tabel V- 17 Hasil Evaluasi DTABC dengan Bee=30 dan Iterasi=75	V-8
Tabel V- 18 Hasil Evaluasi DTABC dengan Bee=30 dan Iterasi=100	V-8
Tabel V- 19 Hasil Evaluasi DTABC dengan Bee=40 dan Iterasi=10	V-8
Tabel V- 20 Hasil Evaluasi DTABC dengan Bee=40 dan Iterasi=20	V-8
Tabel V- 21 Hasil Evaluasi DTABC dengan Bee=40 dan Iterasi=50	V-9
Tabel V- 22 Hasil Evaluasi DTABC dengan Bee=40 dan Iterasi=75	V-9
Tabel V- 23 Hasil Evaluasi DTABC dengan Bee=40 dan Iterasi=100	V-9
Tabel V- 24 Hasil Evaluasi DTABC dengan Bee=50 dan Iterasi=10	V-9
Tabel V- 25 Hasil Evaluasi DTABC dengan Bee=50 dan Iterasi=20	V-10
Tabel V- 26 Hasil Evaluasi DTABC dengan Bee=50 dan Iterasi=50	V-10
Tabel V- 27 Hasil Evaluasi DTABC dengan Bee=50 dan Iterasi=75	V-10
Tabel V- 28 Hasil Evaluasi DTABC dengan Bee=50 dan Iterasi=100	V-10
Tabel V- 29 Hasil Perbandingan Nilai Akurasi	V-11
Tabel V- 30 Hasil Akhir Akurasi <i>Decision Tree</i> dan DTABC	V-12

DAFTAR GAMBAR

	Halaman
Gambar II- 1 Struktur <i>Decision Tree</i>	II-5
Gambar II- 2 Diagram Algoritma <i>artificial Bee Colony</i>	II-8
Gambar II- 3 <i>Rational Unified Process</i> (RUP)	II-14
Gambar III- 1 Diagram Tahapan Penelitian	III-4
Gambar III- 2 Kerangka Kerja Penelitian	III-6
Gambar IV- 1 Diagram <i>Use Case</i>	IV-5
Gambar IV- 2 Rancangan Antarmuka menu Model Information	IV-10
Gambar IV- 3 Rancangan Antarmuka Menu Prediction	IV-11
Gambar IV- 4 Rancangan Antarmuka Menu Evaluation	IV-12
Gambar IV- 5 Diagram Aktivasi Klasifikasi <i>Decision Tree</i> C4.5 dengan <i>Artificial Bee Colony</i>	IV-14
Gambar IV- 6 Diagram Aktivasi Memilih File	IV-14
Gambar IV- 7 Diagram Aktivasi Input Data Baru dan Memilih Metode untuk Prediksi	IV-15
Gambar IV- 8 Diagram <i>Sequence</i> Memilih File	IV-15
Gambar IV- 9 Diagram <i>Sequence</i> Input Data Baru dan Memilih Metode untuk Prediksi	IV-16
Gambar IV- 10 Diagram Kelas	IV-17
Gambar IV- 11 Antarmuka Menu Model Information	IV-19
Gambar IV- 12 Antarmuka Menu Evaluation	IV-19
Gambar IV- 13 Antarmuka Menu Prediction	IV-20
Gambar V- 1 Grafik Intensitas Persebaran Fitur Seleksi Fitur ABC	IV-12
Gambar V- 2 Grafik Perbandingan Waktu Komputasi	IV-13

DAFTAR LAMPIRAN

	Halaman
Lampiran 1. Dataset.....	xxiii
Lampiran 2. Analisis <i>Decision Tree</i> C4.5	xxiii
Lampiran 3. Analisis <i>Artificial Bee Colony</i>	xxiii
Lampiran 4. Kode Program	xxiii

BAB I

PENDAHULUAN

1.1 Pendahuluan

Pada bab ini membahas mengenai latar belakang masalah, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, sistematika penulisan dan kesimpulan judul yang diangkat secara rinci. Bab ini menjelaskan tentang penjelasan secara umum terhadap penelitian yang dikerjakan.

1.2 Latar Belakang Masalah

Secara umum, diabetes dapat menyerang siapa saja, dalam kondisi apa saja, dan dapat berlangsung seumur hidup. Faktanya, sebagian besar gangguan diabetes tidak terdiagnosis dengan benar, yang dapat mengakibatkan komplikasi fatal dengan penyakit ini (Santoso, Syukur, dan Fanani 2018).

Melakukan diagnosis pada diabetes merupakan salah satu usaha untuk mengatasi masalah ini. Diharapkan bahwa identifikasi ini akan mengingatkan individu dengan diabetes terhadap kemungkinan kesulitan di kemudian hari. Diabetes adalah suatu kondisi yang dicatat secara rutin untuk memeriksa pasien dan melakukan tindakan pencegahan. Sistem klasifikasi penyakit diabetes menjadi sangat penting dalam dunia medis. Dengan adanya sistem ini, diagnosis diabetes dapat dilakukan dengan lebih akurat dan efisien, memungkinkan para profesional kesehatan untuk memberikan perawatan yang lebih tepat sasaran.

Pada penelitian (Hana, 2020) menerapkan Algoritma C4.5 untuk melakukan klasifikasi apakah seseorang terkena penyakit diabetes atau tidak

mengidap penyakit tersebut. Hasil Pengujian menghasilkan akurasi yang cukup besar yaitu 97,12 % *Precision* sebesar 93,02% %, dan *Recall* sebesar 100,00%. Adapun Kurva ROC (*Receiver Operating Characteristic*) menunjukkan algoritma C4.5 memiliki nilai AUC sebesar 0.994 yang artinya *Excellent Classification*, ini menunjukkan bahwa menerapkan Algoritma C4.5 untuk melakukan klasifikasi apakah seseorang terkena penyakit diabetes atau tidak mengidap penyakit tersebut.

Decision Tree memiliki kelemahan jika jumlah kelas dan atribut berjumlah banyak sehingga dapat menimbulkan *overlap* dan salah satu solusinya adalah menggunakan *Random Forest*, namun *Random Forest* juga bisa kurang optimal jika menggunakan *dataset* yang tidak seimbang (Pangastuti, 2018).

Salah satu cara untuk menghindari *overlap* yaitu dengan seleksi fitur. Hal ini dapat meningkatkan akurasi model karena *Decision Tree* akan lebih efisien dalam membagi data berdasarkan fitur-fitur yang penting. Seleksi fitur memiliki beberapa manfaat, termasuk pengurangan *overfitting* yang dapat menghilangkan *noise* dan data yang berlebihan. Selain itu, karena seleksi fitur dapat menghilangkan fitur yang tidak terlalu berpengaruh terhadap kategori, maka akurasi dapat ditingkatkan. Selain itu, waktu komputasi berkurang karena algoritma pembelajaran dapat beroperasi lebih cepat dan efektif jika dimensi dataset lebih sederhana.

Dalam sebuah penelitian yang dilakukan (Junita dan Bachtiar, 2020) melakukan penelitian untuk klasifikasi Aktivitas Manusia menggunakan Algoritme *Decision Tree* C4.5 dan *Information Gain* untuk Seleksi Fitur, penelitian dilakukan karena upaya sebelumnya yang dilakukan oleh penelitian lain

menghasilkan hasil yang kurang ideal. Selain itu, dataset fitur dimensi tinggi yang digunakan untuk pengenalan aktivitas manusia memiliki masalah yang menyebabkan waktu komputasi yang lama dan *overfitting* dalam model klasifikasi.

Berdasarkan hasil pengujian yang telah dilakukan menggunakan 90% dari total fitur (561) dan 10 untuk *depth tree* maksimum menghasilkan akurasi sebesar 81%. Dalam investigasi ini *depth tree* yang optimal untuk digunakan adalah 10. Hasil akurasi ini berlaku ketika menguji dengan *depth tree* maksimum 15 dan 20 dan waktu komputasi yang lebih lama. Hal ini menunjukkan bahwa membatasi *depth tree* dapat meningkatkan akurasi klasifikasi.

Setelah melakukan seleksi fitur, agar mempermudah dalam seleksinya dapat menggunakan algoritma optimasi. Di antara metode yang dapat diterapkan adalah algoritma *Artificial Bee Colony*. Hal ini merupakan hasil dari kemampuan algoritma untuk mengidentifikasi makanan terbaik untuk masalah pengoptimalan (Pradnyana et.al., 2018). Baykasoglu (2007) menyatakan bahwa algoritma-algoritma yang terinspirasi oleh lebah memiliki potensi yang sangat menjanjikan untuk pemodelan dan pemecahan masalah optimasi yang kompleks.

Oleh karena itu, penelitian ini bertujuan untuk mengoptimalkan algoritma *Decision Tree* menggunakan metode *Artificial Bee Colony* untuk meningkatkan akurasi klasifikasi data penderita penyakit diabetes. Dengan demikian, penelitian ini diharapkan dapat memberikan kontribusi dalam meningkatkan kinerja algoritma *Decision Tree* dalam klasifikasi data penderita penyakit diabetes, sehingga dapat membantu dokter dalam membuat keputusan dalam pengobatan pasien diabetes.

Dalam penelitian ini, sebuah perangkat lunak dikembangkan untuk mengoptimalkan klasifikasi data penyakit diabetes dengan mempertimbangkan masalah yang telah disebutkan sebelumnya. Adapun metode yang digunakan adalah *Decision Tree*. Kemudian, penulis membuat penelitian dengan judul “Optimasi Algoritma *Decision Tree* Menggunakan *Artificial Bee Colony* Untuk Klasifikasi Data Penderita Penyakit Diabetes”. Untuk mengukur kinerja tingkat kesesuaian klasifikasi penelitian ini dibatasi dengan adanya presentasi hasil akurasi.

1.3 Rumusan Masalah

Beberapa rumusan masalah berdasarkan latar belakang yang telah disebutkan yang akan dibahas dalam penelitian ini adalah:

1. Apakah *Artificial Bee Colony* dapat menyeleksi fitur pada *Decision Tree* untuk klasifikasi data pada penyakit diabetes?
2. Apakah seleksi fitur pada *Decision Tree* dengan *Artificial Bee Colony* dapat memengaruhi waktu komputasi?
3. Bagaimana hasil akurasi algoritma *Decision Tree* dalam klasifikasi data penderita penyakit diabetes sebelum dan setelah dioptimalkan menggunakan *Artificial Bee Colony*?

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah, tujuan penelitian yang dilakukan adalah sebagai berikut:

1. Membangun sistem klasifikasi penyakit diabetes dengan melakukan optimasi seleksi fitur pada *Decision Tree* menggunakan *Artificial Bee Colony*.
2. Mengetahui perbandingan waktu komputasi algoritma *Decision Tree* dalam klasifikasi data penderita penyakit diabetes sebelum dan setelah dioptimalkan menggunakan *Artificial Bee Colony*.
3. Mengetahui hasil akurasi algoritma *Decision Tree* dalam klasifikasi data penderita penyakit diabetes sebelum dan setelah dioptimalkan menggunakan *Artificial Bee Colony*.

1.5 Manfaat Penelitian

Diharapkan pada penelitian yang dijalankan dapat memberi manfaat sebagai berikut:

1. Hasil penelitian ini diharapkan dapat membantu pengguna untuk mengetahui perkiraan terkena penyakit diabetes.
2. Menambah literatur dan kontribusi dalam bidang data mining dan machine learning dalam klasifikasi penyakit diabetes, khususnya dalam penggunaan algoritma optimasi *Artificial Bee Colony* pada algoritma klasifikasi *Decision Tree*.
3. Hasil penelitian ini dapat digunakan sebagai referensi untuk melakukan prediksi penyakit diabetes dan sebagai landasan untuk penelitian selanjutnya.

1.6 Batasan Masalah

Agar ruang lingkup masalah pada penelitian ini jelas dan terhindar dari adanya penyimpangan, adapun batasan masalah pada penelitian ini yaitu data yang digunakan adalah data sekunder yang diambil dari situs <https://www.kaggle.com> yaitu *Diabetes Dataset* yang diperbarui pada tahun 2022. Kumpulan data ini berasal dari *National Institute of Diabetes and Digestive and Kidney Diseases*.

1.7 Sistematika Penulisan

Sistematika penulisan skripsi ini adalah sebagai berikut:

BAB I. PENDAHULUAN

Pada bab ini membahas mengenai latar belakang masalah, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, sistematika penulisan dan kesimpulan dari judul yang diangkat.

BAB II. KAJIAN LITERATUR

Pada bab ini membahas tentang dasar-dasar teori yang digunakan dalam penelitian yaitu teori klasifikasi data, optimasi, algoritma *Decision Tree*, *Artificial Bee Colony*, serta penelitian-penelitian lain yang relevan.

BAB III. METODOLOGI PENELITIAN

Bab ini akan membahas tahapan yang akan dilaksanakan pada penelitian ini. Masing-masing rencana tahapan penelitian dideskripsikan dengan rinci dengan mengacu pada suatu kerangka kerja. Di akhir bab ini berisi perancangan manajemen proyek pada pelaksanaan penelitian.

BAB IV. PENGEMBANGAN PERANGKAT LUNAK

Bab ini akan membahas proses pengembangan perangkat lunak menggunakan algoritma *Decision Tree* dan *Artificial Bee Colony*. Implementasi dari tiap fase berdasarkan metode *Rational Unified Process*.

BAB V. HASIL DAN ANALISIS PENELITIAN

Pada bab ini akan dijelaskan mengenai proses pengujian berdasarkan format yang sudah dirancangkan dari perangkat lunak serta analisis terhadap hasil penelitian.

BAB VI. KESIMPULAN DAN SARAN

Pada bab ini akan dijelaskan mengenai kesimpulan dari penelitian dan pengujian yang dilakukan beserta saran yang diharapkan untuk digunakan sebagai referensi dalam pengembangan penelitian selanjutnya.

1.8 Kesimpulan

Berdasarkan rumusan masalah yang diteliti, tujuan penelitian, dan batasan masalah yang diteliti dapat ditetapkan berdasarkan latar belakang masalah. Judul penelitian yang dilakukan adalah “Optimasi Algoritma *Decision Tree* Menggunakan *Artificial Bee Colony* untuk Klasifikasi Data Penderita Penyakit Diabetes”.

DAFTAR PUSTAKA

- Allugunti, V. R., Kishor Kumar Reddy, C., Elango, N. M., & Anisha, P. R. (2020). *Prediction of diabetes using internet of things (iot) and decision trees: SLDPS*. In *Advances in Intelligent Systems and Computing* (pp. 453–461). Springer Singapore. http://dx.doi.org/10.1007/978-981-15-5679-1_43
- Apriliah, W., Kurniawan, I., Baydhowi, M., & Haryati, T. (2021). SISTEMASI: Jurnal Sistem Informasi Prediksi Kemungkinan Diabetes pada Tahap Awal Menggunakan Algoritma Klasifikasi Random Forest. *Jurnal Sistem Informasi, 10*(1), 163–171. <http://sistemasi.ftik.unisi.ac.id>
- Azhari, M., Situmorang, Z., & Rosnelly, R. (2021). Perbandingan Akurasi, Recall, dan Presisi Klasifikasi pada Algoritma C4.5, Random Forest, SVM dan Naive Bayes. *Jurnal Media Informatika Budidarma, 5*(2), 640. <https://doi.org/10.30865/mib.v5i2.2937>
- Baykasoğlu, A., Özbakır, L., & Tapkan, P. (2007). *Artificial Bee Colony Algorithm and Its Application to Generalized Assignment Problem*. *Swarm Intelligence, Focus on Ant and Particle Swarm Optimization*, Felix T. S. Chan and Manoj Kumar Tiwari (Ed.), ISBN: 978-3-902613-09-7
- Butt, U. M., Letchmunan, S., Ali, M., Hassan, F. H., Baqir, A., & Sherazi, H. H. R. (2021). Machine learning based diabetes classification and prediction for healthcare applications. *Journal of Healthcare Engineering, 2021*, 1–17. <https://doi.org/10.1155/2021/9930985>
- Chandrasekar, P., Qian, K., Shahriar, H., & Bhattacharya, P. (2017). *Improving the Prediction Accuracy of Decision Tree Mining with Data Preprocessing*. *Proceedings - International Computer Software and Applications Conference, 2*, 481–484. <https://doi.org/10.1109/COMPSAC.2017.146>

- Hana, F. M. (2020). Klasifikasi Penderita Penyakit Diabetes Menggunakan Algoritma *Decision Tree C4.5*. *Jurnal Sistem Komputer Dan Kecerdasan Buatan (Siskom-KB)*, 4(1), 32–39.
- Hadianto, N., Novitasari, H. B., & Rahmawati, A. (2019). Klasifikasi Peminjaman Nasabah Bank Menggunakan Metode Neural Network. *Jurnal Pilar Nusa Mandiri*, 15(2), 163–170. <https://doi.org/10.33480/pilar.v15i2.658>
- Hapsari, R. K., Utoyo, M. I., Rulaningtyas, R., & Suprajitno, H. (2020). *Iris segmentation using Hough Transform method and Fuzzy C-Means method*. *Journal of Physics: Conference Series*, 1477(2). <https://doi.org/10.1088/1742-6596/1477/2/022037>
- Hutahaean, D. J., Wardani, N. H., & Purnomo, W. (2019). Pengembangan Sistem Informasi Penyewaan Gedung Berbasis Web dengan Metode Rational Unified Process (RUP) (Studi Kasus: Wisma Rata Medan). *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 3(6), 5789–5798. <http://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/5578>
- Jiang, W., Liu, G., Zhao, X., & Yang, F. (2019). *Cross-Subject Emotion Recognition with a Decision Tree Classifier Based on Sequential Backward Selection*. *Proceedings - 2019 11th International Conference on Intelligent Human-Machine Systems and Cybernetics, IHMSC 2019*, 1, 309–313. <https://doi.org/10.1109/IHMSC.2019.00078>
- Junita, V., & Bachtiar, F. A. (2020). Klasifikasi Aktivitas Manusia menggunakan Algoritma *Decision Tree C4.5* dan *Information Gain* untuk Seleksi Fitur. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 3(10), 9426–9433. <http://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/6446>
- Kurniawan, F., & Ivandari. (2017). Komparasi Algoritma Data Mining Untuk Klasifikasi Penyakit Kanker Payudara. *IC-Tech*, XII(1), 1–8. <http://jurnal.stmik-wp.ac.id>

- Lutfi, M. (2018). Klasifikasi Kualitas Mutu Jagung Dengan Menggunakan *Metode Decision Tree* Pada Dinas Pertanian Bojonegoro. (2018). *Agromix*, 9(1). <https://doi.org/10.35891/agx.v9i1.1360>
- Nasrullah, A. H. (2021). Implementasi Algoritma Decision Tree Untuk Klasifikasi Produk Laris. *Jurnal Ilmiah Ilmu Komputer*, 7(2), 45–51. <https://doi.org/10.35329/jiik.v7i2.203>
- Pangastuti, S. S. (2018). *Perbandingan Metode Ensemble Random Forest Dengan Smote-Boosting Dan Smote-Bagging Pada Klasifikasi Data Mining Untuk Kelas Imbalance*. Surabaya: Tesis Institut Teknologi Sepuluh Nopember
- Putra, P. D., Sukemi, S., & Rini, D. P. (2021). Peningkatan Akurasi Klasifikasi Backpropagation Menggunakan Artificial Bee Colony dan K-NN Pada Penyakit Jantung. *Jurnal Media Informatika Budidarma*, 5(1), 208. <https://doi.org/10.30865/mib.v5i1.2634>
- Hermawan, M, A., Hidayat, N., Setiawan, B, D. 2017. *Sistem Optimasi Rute Tempat Wisata Kuliner DI Malang Menggunakan Algoritma Bee Colony*. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 1(3):215-223.
- Saifullah, S., Zarlis, M., Zakaria, Z., & Sembiring, R. W. (2017). Analisa Terhadap Perbandingan Algoritma Decision Tree Dengan Algoritma Random Tree Untuk Pre-Processing Data. *J-SAKTI (Jurnal Sains Komputer Dan Informatika)*, 1(2), 180. <https://doi.org/10.30645/j-sakti.v1i2.41>
- Santoso, F., Syukur, A., & Fanani, A. Z. (2018). Algoritma C4.5 Dengan Particle Swarm Optimization Untuk Klasifikasi Lama Menghafal Al-Quran Pada Santri Mahadul Quran. *Jurnal Teknologi Informasi*, 14, 92–103.

- Setio, P. B. N., Saputro, D. R. S., & Bowo Winarno. (2020). Klasifikasi Dengan Pohon Keputusan Berbasis Algoritme C4.5. *PRISMA, Prosiding Seminar Nasional Matematika*, 3, 64–71.
- Sugioko, A. (2013). Perbandingan Algoritma Bee Colony dengan Algoritma Bee Colony Tabu List dalam Penjadwalan Flow Shop. *Jurnal METRIS*, 14(02), 113–120.
- Viloria, A., Herazo-Beltran, Y., Cabrera, D., & Pineda, O. B. (2020). Diabetes Diagnostic Prediction Using Vector Support Machines. *Procedia Computer Science*, 170, 376–381. <https://doi.org/10.1016/j.procs.2020.03.065>
- Wei, X., Ma, L., Zhang, H., & Liu, Y. (2021). Multi-core-, multi-thread-based optimization algorithm for large-scale traveling salesman problem. *Alexandria Engineering Journal*, 60(1), 189–197. <https://doi.org/10.1016/j.aej.2020.06.055>
- Widiyati, D. K., Wati, M., & Pakpahan, H. S. (2018). Penerapan Algoritma ID3 Decision Tree Pada Penentuan Penerima Program Bantuan Pemerintah Daerah di Kabupaten Kutai Kartanegara. *Jurnal Rekayasa Teknologi Informasi (JURTI)*, 2(2), 125. <https://doi.org/10.30872/jurti.v2i2.1864>
- Zhang, Q., Bu, X., Gao, H., Li, T., & Zhang, H. (2023). A hierarchical learning based artificial bee colony algorithm for numerical global optimization and its applications. *Applied Intelligence*, 54(1), 169–200. <https://doi.org/10.1007/s10489-023-05202-2>