

**KLASIFIKASI PENYAKIT GINJAL MENGGUNAKAN
ALGORITMA *ARTIFICIAL BEE COLONY* DAN
*GENETIC MODIFIED KNN***



**OLEH:
ARDINA ARIANI
09042621721006**

**PROGRAM STUDI MAGISTER ILMU KOMPUTER
FAKULTAS ILMU KOMPUTER
UNIVERSITAS SRIWIJAYA
2020**

LEMBAR PENGESAHAN

**KLASIFIKASI PENYAKIT GINJAL MENGGUNAKAN
ALGORITMA *ARTIFICIAL BEE COLONY* DAN
*GENETIC MODIFIED KNN***

TESIS

Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Magister

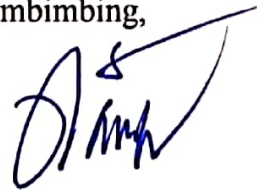
OLEH :
ARDINA ARIANI
09042621721006

Mengetahui,
Koordinator Program Studi
Magister Ilmu Komputer



Dr. Ir. Sukemi, M.T.
NIP 196612032006041001

Palembang, September 2020
Pembimbing,



Samsuryadi, M.Kom., Ph.D.
NIP 197102041997021003

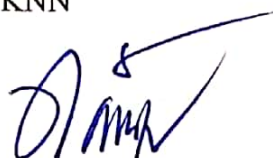
HALAMAN PERSETUJUAN

Pada hari Rabu tanggal 30 Juli 2020 telah dilaksanakan ujian Tesis II secara daring oleh Program Studi Magister Ilmu Komputer Fakultas Ilmu Komputer Universitas Sriwijaya.

Nama : Ardina Ariani
NIM : 09042621721006
Judul : Klasifikasi Penyakit Ginjal Menggunakan Algoritma *Artificial Bee Colony* dan *Genetic Modified KNN*

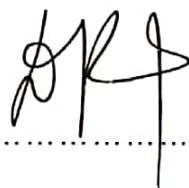
1. Pembimbing

Samsuryadi, M.Kom., Ph.D.
NIP. 197102041997021003



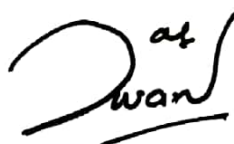
2. Penguji I

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




3. Penguji II

Dr. Iwan Pahendra A.S., M.T.
NIP. 197403222002121002



Palembang, September 2020
Mengetahui,
Koordinator Program Studi Magister Ilmu Komputer



Dr. Ir. Sukemi, M.T.
NIP 196612032006041001

LEMBAR PERNYATAAN

Yang bertanda tangan di bawah ini :

Nama : Ardina Ariani
NIM : 09042621721006
Program Studi : Magister Ilmu Komputer
Judul Tesis : Klasifikasi Penyakit Ginjal Menggunakan Algoritma *Artificial Bee Colony* (ABC) dan *Genetic Modified K-Nearest Neighbor* (GMKNN)

Hasil Pengecekan Software iThenticate/Turnitin : 12 %

Menyatakan bahwa laporan tesis saya merupakan hasil karya sendiri dan bukan hasil penjiplakan/plagiat. Apabila ditemukan unsur penjiplakan/plagiat dalam laporan tesis 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, Agustus 2020



Ardina Ariani
NIM. 09042621721006

KATA PENGANTAR

Puji syukur penulis panjatkan kepada Allah SWT atas limpahan karunia, rahmat dan pertolongan-Nya sehingga penulis dapat menyelesaikan Tesis dengan judul “Klasifikasi Penyakit Ginjal Menggunakan Algoritma *Artificial Bee Colony* (ABC) dan *Genetic Modified K-Nearest Neighbor* (GMKNN)”. Tesis ini diajukan untuk melengkapi salah satu syarat memperoleh gelar Magister di program studi Magister Ilmu Komputer Universitas Sriwijaya.

Penulis berharap Tesis ini dapat bermanfaat bagi orang banyak, meski masih banyak kekurangan dan jauh dari kesempurnaan. Dalam penyusunan Tesis ini, penulis banyak mendapat bimbingan, dukungan dan bantuan dari berbagai pihak, baik moril maupun materil, sehingga Tesis ini dapat diselesaikan. Dengan ketulusan hati, penulis mengucapkan terimakasih kepada:

1. Orang tua, Papa Amzoni dan Mamak Salamah yang selalu memberikan doa, restu, dukungan dan ridho yang diberikan kepada penulis,
2. Saudara kandung penulis, Werosa Okti Franiko, Aris Eto Hendra dan Vonny Oktari yang memberikan semangat yang positif terhadap penulis,
3. Bapak Samsuryadi, M.Kom., Ph.D. selaku Pembimbing Tesis yang selalu mengarahkan, memberi nasihat dan dukungan selama pembuatan Tesis dan publikasi-publikasi sebagai syarat memperoleh gelar Magister.
4. Dr. Ir. Sukemi, M.T selaku Koordinator Program Studi Magister Ilmu Komputer Fakultas Ilmu Komputer Universitas Sriwijaya sekaligus sebagai atasan dalam bekerja yang selalu memonitor dan memberikan dukungan dalam menyelesaikan studi ini,
5. Ibu Dian Palupi Rini, M.Kom., Ph.D. selaku Penguji I dan Bapak Dr. Iwan Pahendra, M.T selaku Penguji II yang sejak awal Sidang Proposal, Seminar Hasil, dan Sidang Tesis II memberikan masukan berupa arahan dan saran untuk Tesis yang lebih baik,
6. Dekan beserta jajaran pimpinan Fakultas Ilmu Komputer Universitas Sriwijaya,
7. Sahabat-sahabat BLU 2015 dan Teman-teman Mahasiswa Program Studi Magister Ilmu Komputer yang tidak dapat penulis sebutkan satu per satu,
8. Bapak/Ibu dosen Program Studi Magister Ilmu Komputer Universitas Sriwijaya.

Penulis yakin segalanya adalah skenario yang diberikan Allah SWT, sehingga penulis berterima kasih kepada segenap orang yang terlibat dalam drama penyelesaian studi ini, sedih ataupun bahagia tetaplah penulis syukuri adanya.

Akhir kata, dengan segala kerendahan hati dan keterbatasan, penulis berharap Tesis ini menghasilkan sesuatu yang bermanfaat, khususnya bagi Fakultas Ilmu Komputer Universitas Sriwijaya secara langsung ataupun tidak langsung sebagai sumbangan pikiran dalam peningkatan mutu pembelajaran dan penelitian.

Palembang, Agustus 2020

Penulis,
Ardina Ariani

CLASSIFICATION OF KIDNEY DISEASE USING ARTIFICIAL BEE COLONY ALGORITHM AND GENETIC MODIFIED KNN

Ardina Ariani

Abstract

The current development of artificial intelligence systems makes health care systems helpful in detecting disease. Early detection of kidney disease is important to recognize symptoms to prevent further damage. This study introduces a kidney disease classification system using the Artificial Bee Colony (ABC) algorithm and Genetic Modified K-Nearest Neighbor (GMKNN), hereinafter referred to as GMKNN-ABC. The ABC algorithm is used as feature selection to determine the symptoms that affect kidney disease with KNN modified with the Genetic Algorithm (AG) used in the classification stage. This research has 3 (three) stages, namely: pre-processing, feature selection, and classification. The performance of the proposed GMKNN-ABC method was compared with the MKNN method without feature selection using ABC, and the MKNN method with feature selection using ABC (MKNN-ABC). The average accuracy rate of the classification results of kidney disease using GMKNN-ABC was 98.27% for 120 test data. The GMKNN-ABC method provides the highest results compared to the results of classification testing using the MKNN, MKNN-ABC methods and other studies with the same dataset.

Keyword: Classification, Kidney Disease, Feature Selection, Modified K-Nearest Neighbor, Genetic Algorithm, Artificial Bee Colony

ABSTRAK

Perkembangan sistem kecerdasan buatan saat ini membuat sistem perawatan kesehatan terbantu dalam mendeteksi penyakit. Deteksi dini penyakit ginjal penting dilakukan untuk mengenali gejala supaya tercegah kerusakan yang lebih parah. Penelitian ini memperkenalkan sistem klasifikasi penyakit ginjal menggunakan algoritma *Artificial Bee Colony* (ABC) dan *Genetic Modified K-Nearest Neighbor* (GMKNN) yang selanjutnya disebut GMKNN-ABC. Algoritma ABC digunakan sebagai seleksi fitur untuk menentukan gejala-gejala yang mempengaruhi penyakit ginjal dengan KNN yang dimodifikasi dengan Algoritma Genetik (AG) digunakan pada tahap klasifikasi. Penelitian ini memiliki 3 (tiga) tahap, yaitu: pra-pemrosesan, pemilihan fitur, dan klasifikasi. Performansi dari metode GMKNN-ABC yang diusulkan dibandingkan dengan metode MKNN tanpa seleksi fitur menggunakan ABC, dan metode MKNN dengan seleksi fitur menggunakan ABC (MKNN-ABC). Tingkat akurasi rata-rata dari hasil klasifikasi penyakit ginjal menggunakan GMKNN-ABC sebesar 98,27% untuk 120 data uji. Metode GMKNN-ABC memberikan hasil tertinggi dibandingkan dengan hasil pengujian klasifikasi menggunakan metode MKNN, MKNN-ABC dan penelitian lainnya dengan dataset yang sama.

Kata Kunci: Klasifikasi, Penyakit Ginjal, Seleksi Fitur, *Modified K-Nearest Neighbor*, Algoritma Genetika, *Artificial Bee Colony*

DAFTAR ISI

	Halaman
Halaman Judul	i
Lembar Pengesahan	ii
Halaman Persetujuan	iii
Lembar Pernyataaan	iv
Kata Pengantar	v
Abstrak	vii
Abstract	viii
Daftar Isi	ix
Daftar Gambar	xi
Daftar Tabel	xii
BAB I. PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Perumusan Masalah	3
1.3 Tujuan Penelitian	4
1.4 Manfaat Penelitian	4
1.5 Batasan Masalah	5
1.6 Sistematika Penulisan	5
BAB II. TINJAUAN PUSTAKA	7
2.1 Klasifikasi	7
2.2 Seleksi Fitur (<i>Future Selection</i>)	8
2.3 <i>K-Nearest Neighbor</i>	8
2.4 <i>Modified K-Nearest Neighbor</i>	10
2.4.1 Validasi Data Latih	11
2.4.2 Perhitungan <i>Weight Voting</i>	11
2.5 <i>Genetic Algorithm</i>	12
2.5.1 <i>Selection</i> (Seleksi)	13
2.5.1.1 Seleksi Roda Rolet	14
2.5.2 <i>Crossover</i> (Penyilangan)	14
2.5.2.1 <i>Single Point Crossover</i>	15
2.5.3 <i>Mutation</i> (Mutasi)	15
2.5.3.1 Mutasi <i>Uniform</i>	15

2.6 <i>Artificial Bee Colony</i>	16
2.7 <i>Confusion Matrix</i>	19
2.8 Penelitian Terkait	21
2.8.1 Berdasarkan Data	21
2.8.2 Berdasarkan Metode	22
BAB III. METODOLOGI PENELITIAN	25
3.1 Pendahuluan	25
3.2 Pengumpulan Data	25
3.3 Tahapan Penelitian	26
3.4 <i>Pre-processing</i> data	27
3.5 Pemilihan Fitur Menggunakan <i>Artificial Bee Colony</i> (ABC)	28
3.6 Optimasi parameter k otomatis pada <i>Modified K-Nearest Neighbor</i> (MKNN) Menggunakan <i>Genetic Algorithm</i> (GA)	30
3.7 Perangkat Yang Digunakan pada Penelitian ini	31
3.8 Representasi Data	31
3.9 Rencana Pengujian	33
3.9.1 Pengujian pada hasil klasifikasi dengan MKNN	33
3.9.2 Pengujian pada hasil klasifikasi dengan modifikasi MKNN-GA	34
3.9.3 Pengujian pada hasil klasifikasi dengan seleksi fitur ABC Menggunakan modifikasi MKNN dengan GA	34
3.10 Hasil Penelitian	35
BAB IV. HASIL DAN ANALISIS	36
4.1 Pemakaian Algoritma Klasifikasi	36
4.1.1 Hasil Klasifikasi Menggunakan <i>Modified K-Nearest Neighbor</i>	36
4.1.2 Hasil Klasifikasi Menggunakan <i>Modified K-Nearest Neighbor</i> (MKNN) dan Seleksi Fitur dengan <i>Artificial Bee Colony</i> (ABC)	39
4.1.3 Hasil Klasifikasi Penyakit Ginjal menggunakan Seleksi Fitur menggunakan <i>Artificial Bee Colony</i> (ABC) dan <i>Genetic Modified K-Nearest Neighbor</i> (GMKNN)	42
4.2 Analisis Perbandingan Hasil Klasifikasi	46
4.3 Analisis Perbandingan Penelitian Sebelumnya	50
BAB V KESIMPULAN DAN SARAN	52
5.1 Kesimpulan	52
5.2 Saran	52
Daftar Pustaka	53

DAFTAR GAMBAR

	Halaman
Gambar 2.1 Contoh <i>Single Point Crossover</i>	15
Gambar 2.2 <i>Confusion Matrix</i>	19
Gambar 3.1 Diagram Alir Tahapan Penelitian	27
Gambar 3.2 Diagram Alir <i>Pre-processing</i> data penyakit ginjal kronis	28
Gambar 3.3 Diagram Alir Seleksi Fitur <i>Artificial Bee Colony</i>	29
Gambar 3.4 Optimasi parameter k otomatis pada Modified <i>K-Nearest Neighbor</i> (MKNN) menggunakan <i>Genetic Algorithm</i> (GA)	30
Gambar 4.1 Hubungan Parameter k dengan <i>Confusion Matrix</i> Menggunakan MKNN	39
Gambar 4.2 Hubungan Parameter k dengan <i>Confusion Matrix</i> Menggunakan MKNN-ABC	42
Gambar 4.3 Hubungan Parameter k dengan <i>Confusion Matrix</i> Menggunakan GMKNN-ABC	45
Gambar 4.4 Model hasil klasifikasi penyakit ginjal menggunakan GMKNN-ABC	46
Gambar 4.5 Perbandingan Akurasi 3 metode	47
Gambar 4.6 Perbandingan Presisi 3 metode	48
Gambar 4.7 Perbandingan <i>Recall</i> dan <i>f1-score</i> 3 metode	48
Gambar 4.8 Grafik Perbandingan Hasil Pengujian 3 Metode	49

DAFTAR TABEL

	Halaman
Tabel 3.1 Perangkat Keras yang digunakan	31
Tabel 3.2 Perangkat Lunak yang digunakan	31
Tabel 3.3 Deskripsi Indikasi (Fitur) Penyakit Ginjal	32
Tabel 3.4 Pengujian pada hasil klasifikasi dengan MKNN	33
Tabel 3.5 Pengujian pada hasil klasifikasi dengan MKNN-GA	34
Tabel 3.6 Pengujian pada hasil klasifikasi dengan seleksi fitur ABC Menggunakan GMKNN (GMKNN-ABC)	34
Tabel 4.1 Hasil Klasifikasi Menggunakan MKNN berdasarkan parameter k	37
Tabel 4.2 Pengujian Metode MKNN Berdasarkan <i>Confusion Matrix</i>	38
Tabel 4.3 Hasil Klasifikasi Menggunakan MKNN-ABC berdasarkan parameter k	40
Tabel 4.4 Pengujian Metode MKNN-ABC Berdasarkan <i>Confusion Matrix</i>	41
Tabel 4.5 Hasil Klasifikasi Menggunakan GMKNN-ABC berdasarkan parameter k	43
Tabel 4.6 Pengujian Metode GMKNN-ABC Berdasarkan <i>Confusion Matrix</i>	44
Tabel 4.7 Perbandingan Hasil <i>Confusion Matrix</i> 3 Metode	46
Tabel 4.7 Perbandingan Penelitian Sebelumnya	51

BAB I. PENDAHULUAN

1.1 Latar Belakang Masalah

Berbagai teknik klasifikasi telah banyak diterapkan dalam bidang kesehatan seperti kanker, parkinson, jantung, liver, diabetes, hepatitis, thalasemia, dan Alzheimer (Khomsah, 2017). Teknik klasifikasi digunakan untuk menemukan faktor-faktor penyebab suatu penyakit dan kelas suatu penyakit berdasarkan gejala-gejala penyakit. Berbagai metode klasifikasi telah diterapkan untuk mengklasifikasi berbagai penyakit. Hasil kinerja metode klasifikasi dapat dipengaruhi oleh data yang mengandung *noise*/derau dan banyaknya fitur.

Semua fitur data tidak selalu memberikan kontribusi yang sama pada hasil akhir. Fitur-fitur utama harus diidentifikasi agar dapat berkontribusi dan digunakan sebagai perwakilan data (Prasartvit dkk, 2012). Fitur-fitur yang tidak relevan dapat mengurangi tingkat akurasi pada proses pengklasifikasian. Untuk mengatasi hal ini, diperlukan pemilihan fitur (*Feature Selection*) supaya menghilangkan fitur-fitur yang tidak relevan untuk meningkatkan kinerja metode yang digunakan pada proses klasifikasi (Schiezaro dan Pedrini, 2013). Pemilihan fitur membantu mempercepat proses klasifikasi dengan mengekstraksi informasi yang relevan dan berguna dari dataset.

Beberapa teknik telah disarankan untuk pemilihan fitur, termasuk pendekatan metaheuristik. Pendekatan metaheuristik merupakan pendekatan yang digunakan untuk menyelesaikan masalah optimasi dengan berorientasi pada pencarian solusi optimal, proses pencarian dilakukan dengan menggunakan beberapa agen yang membentuk sebuah sistem solusi yang berkembang dengan seperangkat aturan atau persamaan matematika selama beberapa iterasi. Iterasi dilakukan hingga solusi ditemukan sesuai dengan kriteria yang ditentukan dan solusi akhir akan dianggap sebagai solusi optimal. Beberapa metode yang termasuk dalam pendekatan metaheuristik seperti *Genetic Algorithm* (GA), *Ant Colony Optimization* (ACO), *Particle Swarm Optimization* (PSO), dan *Artificial*

Bee Colony (ABC). Pendekatan-pendekatan ini telah menghasilkan hasil yang baik dalam memperoleh subset fitur yang optimal (Ahmad dkk. 2019). Pada penelitian Zyout dkk. (2011) *Particle Swarm Optimization* (PSO) diterapkan untuk pemilihan fitur pada diagnosis berbasis bentuk dalam *Cluster Microcalcification* dalam Mamografi. Penelitian Wang dkk. (2007) mengimplementasikan *Particle Swarm Optimization* (PSO) sebagai seleksi fitur pada beberapa data UCI, hasilnya menunjukkan bahwa PSO memerlukan waktu eksekusi yang lebih lama dibandingkan eksekusi waktu dengan metode ABC, yaitu pada penelitian Shunmugapriya dan Kanmani, (2017) pada penelitiannya membandingkan kinerja metode *Ant Colony Optimization* (ACO) dan *Artificial Bee Colony* (ABC) sebagai seleksi fitur. Hasil penelitian menunjukkan bahwa metode ABC memberikan kinerja yang lebih baik dibandingkan dengan metode ACO dalam melakukan pengurangan ukuran subset fitur dan tingkat akurasi yang dihasilkan. Setelah fitur terpilih, fitur-fitur akan digunakan dalam proses klasifikasi. Proses klasifikasi bergantung pada metode yang digunakan apakah sesuai dengan dataset yang akan diolah. Beberapa metode klasifikasi telah diterapkan dalam berbagai kasus klasifikasi, seperti Metode *K – Nearest Neighbor* (KNN), Algoritma *K – Means*, dan Jaringan Saraf Tiruan (JST) telah digunakan pada pengklasifikasian penyakit hepatitis. Menurut Behori (2017) metode *K - Nearest Neighbor* (KNN) dinyatakan lebih baik dalam mengklasifikasi penyakit hepatitis dibandingkan dengan metode *Neural Network* (NN) karena menghasilkan nilai RMSE dan MSE yang lebih kecil. Penelitian (Jabbar dkk. 2013), (Kuhkan, 2016) menyebutkan bahwa metode *K - Nearest Neighbor* (KNN) digunakan dalam banyak aplikasi seperti klasifikasi dan interpretasi.

Algoritma KNN telah banyak digunakan sebagai model klasifikasi yang efektif. Data yang baru diklasifikasikan berdasarkan mayoritas dari kategori K-tetangga terdekat. Jarak antar data dihitung berdasarkan semua atribut data. Banyaknya atribut mempengaruhi hasil penghitungan jarak antar data dan akan mempengaruhi tingkat keakuratan. KNN memiliki beberapa kelebihan, diantaranya adalah pelatihan sangat cepat, sederhana dan mudah dipelajari, tahan terhadap data pelatihan yang memiliki derau, dan efektif jika data pelatihan besar. Namun, KNN juga memiliki kekurangan yaitu nilai k yang bias, komputasi

kompleks, keterbatasan memori, dan mudah tertipu dengan atribut yang tidak relevan. Nilai k memiliki dampak besar pada kinerja klasifikasi terutama dengan data dimensi tinggi. Ketergantungan terhadap nilai k terbaik, memerlukan penelitian lebih lanjut untuk meningkatkan akurasi dengan nilai k yang baik (Jabbar dkk. 2013). Penelitian Kuhkan (2016) mengoptimasi teknik klasifikasi KNN menghasilkan tingkat akurasi lebih tinggi dibandingkan dengan metode lain seperti J48 dan *Naïve Bayes*. Penelitian lain menggunakan *Genetic Algorithm* (GA) untuk mencari nilai optimal dari parameter k pada metode pengembangan KNN yaitu metode *Modified KNN* (MKNN) yang memperoleh hasil yang lebih baik dalam hal tingkat akurasi klasifikasi dan mendapatkan parameter k optimal secara otomatis (Mutrofin dkk. 2014). Penelitian Mutrofin dkk. (2014) memperlihatkan bahwa metode GA dapat digunakan untuk mendapatkan parameter k optimal secara otomatis dimana memperbaiki kelemahan metode KNN dalam penentuan parameter k . Sedangkan metode KNN diperlukan peningkatan kinerja dengan *Modified KNN* (MKNN) dimana terdapat perhitungan nilai validitas untuk mempertimbangkan validitas antar data latih agar dapat mengetahui nilai stabilitas dan kekokohan sampel data pelatihan pada ruang fitur yang selanjutnya dilakukan *voting* untuk menentukan kelas berdasarkan bobot data yang dihasilkan. Adanya perhitungan nilai validitas dan *weighted voting* pada metode MKNN dapat menaikkan tingkat akurasi yang dihasilkan dibandingkan dengan KNN tradisional (Ravi dkk. 2019).

Berdasarkan pada penjelasan tersebut, maka tesis ini akan menerapkan pemilihan fitur menggunakan *Artificial Bee Colony* (ABC), pencarian parameter k secara otomatis menggunakan *Genetic Algorithm* (GA) pada *Modified K - Nearest Neighbor* (*Genetic Modified KNN*) yang selanjutnya disebut GMKNN-ABC. Kombinasi metode ini akan dipakai untuk melihat tingkat akurasi yang dihasilkan pada klasifikasi penyakit ginjal.

1.2 Perumusan Masalah

Jumlah fitur yang banyak mempengaruhi tingkat keakuratan suatu proses klasifikasi. Sehingga dibutuhkan pemilihan fitur untuk meningkatkan akurasi

terhadap pengklasifikasian. Selain itu, penentuan parameter k juga berpengaruh terhadap proses dan hasil klasifikasi. Berdasarkan kenyataan tersebut, disusunlah pertanyaan penelitian ini sebagai berikut:

1. Bagaimana mengembangkan kerangka kerja pemilihan fitur pada penyakit ginjal dengan *Artificial Bee Colony* (ABC) dan klasifikasi penyakit ginjal dengan metode GMKNN-ABC?
2. Bagaimana pengaruh fitur-fitur yang terpilih terhadap kinerja metode dalam klasifikasi penyakit ginjal?
3. Bagaimana pengaruh parameter k pada kinerja metode *Genetic Modified K - Nearest Neighbor* (MKNN) untuk mengklasifikasi penyakit ginjal?

1.3 Tujuan Penelitian

Tujuan penelitian ini adalah sebagai berikut:

1. Mengembangkan kerangka kerja pemilihan fitur dengan *Artificial Bee Colony* (ABC) penyakit ginjal.
2. Menganalisis pengaruh pemilihan fitur menggunakan *Artificial Bee Colony* (ABC) dan mengukur hasil klasifikasi dengan *Modified K - Nearest Neighbor* (MKNN) - *Genetic Algorithm* (GA).
3. Menganalisis pengaruh parameter k terhadap kinerja metode *Genetic Algorithm* (GA) pada *Modified K - Nearest Neighbor* (MKNN) untuk melakukan klasifikasi pada penyakit ginjal berdasarkan tingkat akurasi, presisi, *recall* dan *f1-score*.

1.4 Manfaat Penelitian

Manfaat penelitian ini adalah :

1. Kerangka kerja penelitian yang dihasilkan dapat menjadi acuan dalam pengklasifikasian penyakit ginjal.
2. Fitur-fitur yang terpilih dapat digunakan untuk menentukan kinerja metode yang dihasilkan.
3. Parameter k terbaik dapat digunakan pada GMKNN untuk melakukan pengklasifikasian.

1.5 Batasan Masalah

Batasan masalah yang akan dibahas pada penelitian ini adalah:

1. Sistem yang dibangun berbentuk simulasi untuk mengklasifikasikan penyakit ginjal.
2. Hasil dari optimasi parameter k dengan GA pada MKNN (GMKNN) dibandingkan dengan hasil yang diperoleh dengan tanpa optimasi parameter k pada MKNN.
3. Hasil penggunaan algoritma ABC sebagai seleksi fitur pada GMKNN dibandingkan dengan hasil yang diperoleh tanpa penggunaan ABC sebagai seleksi fitur pada MKNN.
4. *Dataset* yang digunakan adalah *dataset* publik. *Dataset* yang digunakan adalah *dataset* yang diambil dari *repository* UCI *Machine Learning*, yaitu *chronic kidney disease* (penyakit ginjal kronis).

1.6 Sistematika Penulisan

Penyusunan Tesis ini menggunakan sistematika penulisan sebagai berikut:

1. Bab I Pendahuluan

Bab ini berisi tentang latar belakang masalah, perumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, dan sistematika penulisan.

2. Bab II Tinjauan Pustaka

Bab ini berisi tentang seluruh penjelasan mengenai tinjauan pustaka yang berhubungan dengan permasalahan yang dibahas pada penulisan tesis ini.

3. Bab III Metodologi Penelitian

Bab ini berisi penjelasan secara bertahap dan terperinci tentang langkah-langkah (metodologi) yang digunakan untuk membuat kerangka berfikir dan kerangka kerja dalam menyelesaikan tesis.

4. Bab IV Hasil dan Analisis

Bab ini berisi uraian dan analisis terhadap hasil penelitian, yang disajikan dalam bentuk gambar maupun tabel.

5. Bab V Kesimpulan dan Saran

Bab ini berisi simpulan dari hasil penelitian, apa yang didapatkan dari penelitian ini disampaikan di bab ini. Berdasarkan kesimpulan, didapatkan saran sebagai perencanaan bagi penelitian selanjutnya.

DAFTAR PUSTAKA

- Adeniyi, D. A., Wei, Z., & Yongquan, Y. (2016). Automated web usage data mining and recommendation system using K-Nearest Neighbor (KNN) classification method. *Applied Computing and Informatics*, 12(1), 90–108. <https://doi.org/10.1016/j.aci.2014.10.001>
- Ahmad, S. R., Bakar, A. A., & Yaakub, M. R. (2019). *Ant colony optimization for text feature selection in sentiment analysis*. 23, 133–158. <https://doi.org/10.3233/IDA-173740>
- Akhil Jabbar, M., Deekshatulu, B. L., & Chandra, P. (2013). Classification of Heart Disease Using K- Nearest Neighbor and Genetic Algorithm. *Procedia Technology*, 10, 85–94. <https://doi.org/10.1016/j.protcy.2013.12.340>
- Ayuningrum, N. L. A., & Saptaningtyas, F. Y. (2017). *Implementasi Algoritma Genetika dengan Variasi Crossover dalam Penyelesaian Capacitated Vehicle Routing Problem with Time Windows (CVRPTW) pada Pendistribusian Air Mineral*. 6(3), 62–72.
- Behori, A. (2017). *Komparasi Model Neural Network dan K-NN untuk Mengklasifikasi Kondisi Penderita Penyakit Hepatitis*. 3(2), 82–88.
- Celik, E., Atalay, M., & Kondiloglu, A. (2016). *Intelligent Systems and Applications in Engineering The Diagnosis and Estimate of Chronic Kidney Disease Using the Machine Learning Methods* #. 1–5.
- Chandel, K., Kunwar, V., Sabitha, S., & Choudhury, T. (2017). SPECIAL ISSUE REDSET 2016 OF CSIT A comparative study on thyroid disease detection using K-nearest neighbor and Naive Bayes classification techniques. *CSI Transactions on ICT*. <https://doi.org/10.1007/s40012-016-0100-5>
- Chatterjee, S., Banerjee, S., Basu, P., Debnath, M., & Sen, S. (2017). Cuckoo search coupled artificial neural network in detection of chronic kidney disease. *2017 1st International Conference on Electronics, Materials Engineering and Nano-Technology, IEMENTech 2017*. <https://doi.org/10.1109/IEMENTECH.2017.8077016>
- Fadilla, I., Adikara, P. P., & Perdana, R. S. (2018). Klasifikasi Penyakit Chronic Kidney Disease (CKD) Dengan Menggunakan Metode Extreme Learning

- Machine (ELM). *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 2(10), 3397–3405.
- Gunavathi, C., & Premalatha, K. (2014). *Performance Analysis of Genetic Algorithm with kNN and SVM for Feature Selection in Tumor Classification*. 8(8), 1490–1497.
- Jena, L., & Kamila, N. K. (2015). Distributed Data Mining Classification Algorithms for Prediction of Chronic-Kidney-Disease. *International Journal of Emerging Research in Management & Technology*, 9359(11), 2278–9359. Retrieved from <https://pdfs.semanticscholar.org/a4a1/4013340306524805459a58848e01787aee9b.pdf>
- Khomsah, S. (2017). *Survei pada Penggunaan Teknik Data Mining pada Bidang Kesehatan di Indonesia*. 2017(November), 82–90.
- Kuhkan, M. (2016). *A Method to Improve the Accuracy of K-Nearest Neighbor Algorithm*. 8(6), 90–95.
- Kunwar, V., Chandel, K., Sabitha, A. S., & Bansal, A. (2016). Chronic Kidney Disease Analysis Using Data Mining Classification. 300–305.
- Lestari, M. (2014). Penerapan Algoritma Klasifikasi Nearest Neighbor (K-NN) untuk Mendeteksi Penyakit Jantung. 7 (September 2010), 366–371.
- Luque, A., Carrasco, A., Martín, A., & de las Heras, A. (2019). The impact of class imbalance in classification performance metrics based on the binary confusion matrix. *Pattern Recognition*. <https://doi.org/10.1016/j.patcog.2019.02.023>
- Majeed, P. G., & Kumar, S. (2014). Genetic Algorithms in Intrusion Detection Systems: A Survey. *International Journal of Innovation and Applied Studies ISSN*, 5(3), 2028–9324. Retrieved from <http://www.issr-journals.org/ijias/>
- Masúdia, P. E., & Wardoyo, R. (2012). *Optimasi Cluster Pada Fuzzy C-Means Menggunakan Algoritma Genetika Untuk Menentukan Nilai Akhir 1*. 6(1), 101–110.
- Medjahed, S. A., & Benyettou, A. (2013). *Breast Cancer Diagnosis by using k-Nearest Neighbor with Different Distances and Classification Rules*. (December 2015). <https://doi.org/10.5120/10041-4635>

- Motieghader, H., Najafi, A., Sadeghi, B., & Masoudi-nejad, A. (2017). A hybrid gene selection algorithm for microarray cancer classification using genetic algorithm and learning automata. *Informatics in Medicine Unlocked*. <https://doi.org/10.1016/j.imu.2017.10.004>
- Mutrofin, S., Izzah, A., & Kurniawardhani, A. (2014). *Optimasi Teknik Klasifikasi Modified k Nearest Neighbor Menggunakan Algoritma Genetika*. (January 2016).
- Nikam, S. S. (2015). *ORIENTAL JOURNAL OF A Comparative Study of Classification Techniques in Data Mining Algorithms*.
- Parvin, H., Alizadeh, H., & Minati, B. (2010). A Modification on K-Nearest Neighbor Classifier. *Global Journal of Computer Science and Technology*, 10(14), 37–41.
- Polat, H., Mehr, H. D., & Cetin, A. (2017). *Diagnosis of Chronic Kidney Disease Based on Support Vector Machine by Feature Selection Methods*. <https://doi.org/10.1007/s10916-017-0703-x>
- Prasartvit, T., Banharnsakun, A., Kaewkamnerdpong, B., & Achalakul, T. (2012). Neurocomputing Reducing bioinformatics data dimension with ABC-kNN. *Neurocomputing*, 1–15. <https://doi.org/10.1016/j.neucom.2012.01.045>
- Rady, E. H. A., & Anwar, A. S. (2019). Prediction of kidney disease stages using data mining algorithms. *Informatics in Medicine Unlocked*, 15(March), 100178. <https://doi.org/10.1016/j.imu.2019.100178>
- Ravi, M. R., & Adinugroho, S. (2019). *Implementasi Algoritme Modified K-Nearest Neighbor (MKNN) Untuk Mengidentifikasi Penyakit Gigi Dan Mulut*. 3(3), 2596–2602.
- Saini, I. (2013). QRS detection using K -Nearest Neighbor algorithm (KNN) and evaluation on standard ECG databases. *Journal of Advanced Research*, 4(4), 331–344. <https://doi.org/10.1016/j.jare.2012.05.007>
- Sara, S. B. V. J., & Kalaiselvi, K. (2018). Ensemble swarm behaviour based feature selection and support vector machine classifier for chronic kidney disease prediction. *International Journal of Engineering and Technology (UAE)*, 7(2), 190–195. <https://doi.org/10.14419/ijet.v7i2.31.13438>

- Saxena, K., Khan, Z., & Singh, S. (2014). *Diagnosis of Diabetes Mellitus using K Nearest Neighbor Algorithm*. 2(4), 36–43.
- Schiezaro, M., & Pedrini, H. (2013). Data feature selection based on Artificial Bee Colony algorithm. *EURASIP Journal on Image and Video Processing*, 2013(1), 1–8. <https://doi.org/10.1186/1687-5281-2013-47>
- Shi, Y., Pun, C., Hu, H., & Gao, H. (2016). An Improved Artificial Bee Colony and Its Application. *Knowledge-Based Systems*. <https://doi.org/10.1016/j.knosys.2016.05.052>
- Shunmugapriya, P., & Kanmani, S. (2017). A Hybrid Algorithm Using Ant and Bee Colony Optimization for Feature Selection and Classification (AC-ABC Hybrid). *Swarm and Evolutionary Computation*. <https://doi.org/10.1016/j.swevo.2017.04.002>
- Sinha, P., & Sinha, P. (2015). *Comparative Study of Chronic Kidney Disease Prediction using KNN and SVM*.
- Suryaputra, J., Lubis, C., & Sutrisno, T. (2018). Pemilihan Crossover pada Algoritma Genetika untuk Program Aplikasi Pengenalan Karakter Tulisan Tangan. *Jurnal Ilmu Komputer Dan Sistem Informasi*, 69–72.
- Vijayarani, S., & Dhayanand, S. (2015). Kidney Disease Prediction Using SVM and ANN Algorithms. *International Journal of Computing and Business Research (IJCBR) ISSN (Online)*, 6(2), 2229–6166.
- Wang, X., Yang, J., Teng, X., Xia, W., & Jensen, R. (2007). Feature selection based on rough sets and particle swarm optimization. *Pattern Recognition Letters*, 28(4), 459–471. <https://doi.org/10.1016/j.patrec.2006.09.003>
- Zyout, I., Abdel-qader, I., & Jacobs, C. (2011). *Embedded Feature Selection using PSO-kNN: Shape-Based Diagnosis of Microcalcification Clusters in Mammography* Embedded Feature Selection using PSO-kNN: Shape-Based Diagnosis of Microcalcification Clusters in Mammography. (February 2015), 6–11. <https://doi.org/10.5383/JUSPN.03.01.002>