

**OPTIMASI BOBOT ATRIBUT PADA ALGORITMA C4.5
MENGUNAKAN *PARTICLE SWARM OPTIMIZATION* UNTUK
PREDIKSI GULA DARAH**

TESIS



OLEH :

**DWI MEYLITASARI BR. TARIGAN
09042681620006**

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

LEMBAR PENGESAHAN

OPTIMASI BOBOT ATRIBUT PADA ALGORITMA C4.5 MENGUNAKAN PARTICLE SWARM OPTIMIZATION UNTUK PREDIKSI GULA DARAH

TESIS

Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Magister

OLEH :

DWI MEYLITASARI BR. TARIGAN

09042681620006

Palembang, Oktober 2020

Pembimbing I,



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

Pembimbing II,



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

Mengetahui,
Koordinator Program Studi Magister Ilmu Komputer




Dr. Ir. Sukemi, M.T
196612032006041001

HALAMAN PERSETUJUAN

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

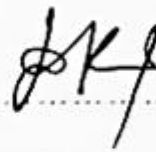
Nama : Dwi Meylitasari Br. Tarigan

NIM : 09042681620006

Judul : Optimasi Bobot Atribut pada Algoritma C4.5 Menggunakan Particle Swarm Optimization untuk Prediksi Gula Darah

1. Pembimbing I

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



2. Pembimbing II

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



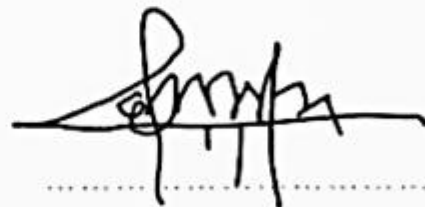
3. Penguji I

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



4. Penguji II

Dr. Ermatita, M.Kom.
NIP. 196709132006042001



Mengetahui,
Koordinator Program Studi
Magister Ilmu Komputer

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

LEMBAR PERNYATAAN

Yang bertanda tangan di bawah ini :

Nama : Dwi Meylitasari Br. Tarigan
NIM : 09042681620006
Judul : Optimasi Bobot Atribut pada Algoritma C4.5 Menggunakan
Particle Swarm Optimization untuk Prediksi Gula Darah

Hasil Pengecekan Software 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 dari siapapun.



Palembang, September 2020



Dwi Meylitasari Br. Tarigan
NIM. 09042681620006

KATA PENGANTAR

Puji syukur atas rahmat dan karunia yang telah Allah SWT berikan sehingga penulis dapat menyelesaikan Tesis ini. Tesis yang berjudul “Optimasi Bobot Atribut pada Algoritma C4.5 Menggunakan Particle Swarm Optimization untuk Prediksi Gula Darah” ini disusun untuk memenuhi salah satu syarat memperoleh gelar Magister pada Program Studi Magister Ilmu Komputer Universitas Sriwijaya.

Pada kesempatan ini, penulis ingin menyampaikan ucapan terima kasih yang tak terhingga kepada pihak-pihak yang telah memberikan dukungan, bimbingan, motivasi dan kemauan kepada penulis untuk menyelesaikan tesis ini, yaitu kepada :

1. Kedua Orang Tua saya serta saudara/I saya yang tercinta, yang tidak henti memberikan motivasi kepada penulis untuk menyelesaikan tesis ini.
2. Bapak Jaidan Jauhari, M.T. selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
3. Bapak Dr. Ir. Sukemi, M.T. selaku Koordinator Program Studi Ilmu Komputer sekaligus dosen penguji II yang telah sabar membimbing penulis.
4. Ibu Dian Palupi Rini, M.Kom., Ph.D. selaku dosen Pembimbing I yang telah banyak membantu dan membimbing penulis dengan sabar hingga dapat menyelesaikan Tesis ini.
5. Bapak Samsuryadi, M.Kom., Ph.D. selaku dosen Pembimbing II yang telah banyak membantu dan membimbing penulis dengan sabar hingga selesai.
6. Ardina Ariani Selaku Admin Prodi Magister Ilmu Komputer atas bantuannya dalam memperlancar kegiatan akademik.
7. Teman-teman magister Ilmu Komputer khususnya angkatan 2016, untuk masa perkuliahan yang menyenangkan.
8. Untuk semua pihak yang telah membantu penyelesaian Tesis ini dan tidak dapat disebutkan satu persatu.

Untuk itu penulis mengharapkan kritik dan saran yang membangun dari semua pihak untuk penyempurnaan tesis ini dan semoga tesis ini dapat bermanfaat bagi pihak yang membutuhkan.

Palembang, Oktober 2020

Penulis

**FEATURE SELECTION IN CLASSIFICATION
OF BLOOD SUGAR DISEASE USING PARTICLE SWARM OPTIMIZATION
(PSO) ON C4.5 ALGORITHM**

Dwi Meylitasari Br. Tarigan

Abstract

Drastic increase of blood sugar level after consuming the certain food because food consumed contain an uncontrolled blood sugar tendence. Many hospital health institution, public health center and clinic which handle the DM patient, from those institution still not provide the data quickly and accurately. The effectively is needed in managing the information with data mining, from large data will generate a new data and provide the information quickly and accurately. In predicting the disease, a lot of the research has been done, the computer science field particularly, with data mining technique to predict the disease using various algorithm such a C45 algorithm. From the research which has been done using C45 algorithm and algorithm C45 with Particle Swarm Optimization (PSO) on set data the effect of physical activity to blood sugar levels in the RSUD H. Abdul Manan Simatupang Kisaran generated the different accuracy. The accuracy value on the tests performed with the C4.5 algorithm is 86%, whereas the accuracy value on the C4.5 algorithm with PSO 95%, so the value difference on its accuracy is 9%. Whereas the evaluation on the ROC curve shows the difference of 0.033.

Keywords : *Datamining, Algorithm C4.5, Particle Swarm Optimization, PSO, classification*

ABSTRAK

Kadar gula darah meningkat dratis setelah mengkonsumsi makanan tertentu karena kecenderungan makanan yang dikonsumsi memiliki kandungan gula darah yang tidak terkontrol. Banyak instansi kesehatan rumah sakit, puskesmas, klinik yang mengatasi berbagai pasien DM, dari beberapa instansi tersebut masih banyak yang belum memberikan data yang cepat dan akurat. Perlu secara efektif dalam mengelola informasi dengan proses data mining, dari data yang besar akan menghasilkan data yang baru dan dapat memberikan informasi secara cepat dan aktual. Dalam melakukan prediksi terhadap penyakit, telah banyak dilakukan penelitian khususnya di bidang computer science, dengan teknik data mining untuk memprediksi penyakit menggunakan berbagai algoritma seperti algoritma C45. Dari penelitian yang telah dilakukan menggunakan algoritma C4.5, dan algoritma C4.5 dengan Particle Swarm Optimization (PSO) pada data set Pengaruh Aktifitas Fisik Terhadap Kadar Gula Darah di RSUD H.Abdul Manan Simatupang Kisaran ini menghasilkan nilai akurasi yang berbeda. Nilai akurasi pada pengujian yang dilakukan dengan algoritma C4.5 sebesar 86%, sedangkan nilai akurasi pada algoritma C4.5 menggunakan PSO sebesar 95%, sehingga selisih nilai pada akurasi nya sebesar 9%. Sedangkan evaluasi pada Kurva ROC menunjukan selisih 0,033.

Kata kunci: data mining, algoritma C4.5, Particle Swarm Optimization, PSO, klasifikasi

DAFTAR ISI

HALAMAN JUDUL	i
LEMBAR PENGESAHAN	ii
HALAMAN PERSETUJUAN	iii
LEMBAR PERNYATAAN	iv
KATA PENGANTAR	v
ABSTRACT	vii
ABSTRAK	viii
DAFTAR ISI	ix
DAFTAR GAMBAR	vi
DAFTAR TABEL	vii

BAB I. PENDAHULUAN

1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	4
1.6 Sistematika Penulisan	5

BAB II. TINJAUAN PUSTAKA

2.1 Penelitian Terkait	6
2.2 Tinjauan Pustaka	7
2.2.1 Diabetes Melitus (DM)	7
2.2.2 Faktor Risiko Diabetes Melitus (DM)	7
2.2.3 <i>Data Mining</i>	8
2.2.4 Fungsi <i>Data Mining</i>	8
2.2.5 Proses <i>Data Mining</i>	9
2.2.6 Klasifikasi	9
2.2.7 Klasifikasi pada Decision Tree	10
2.2.8 Proses Pembentukan pohon keputusan (<i>Decision Tree</i>)	11
2.2.9 Algoritma C4.5	12
2.2.10 Particle Swarm Optimization (PSO)	13
2.2.11 <i>Confusion Matrix</i>	14
2.2.12 <i>Cross Validation</i>	15
2.2.13 Kurva ROC	15

BAB III. METODOLOGI PENELITIAN

3.1 Pengumpulan Data	16
3.2 Pengolahan Data	16
3.2.1 Data Validasi	17

3.2.2	<i>Data Size Reduction</i>	17
3.2.3	Model Algoritma C4.5	17
3.2.4	Evaluasi & Validasi Hasil	17
3.3	Pemodelan	18
3.4	Kerangka Pemikiran	18
3.5	Analisis Algoritma	20
3.5.1	Analisis Pemodela secara umum	20
3.5.2	Analisis Algoritma C4.5	20
3.5.3	Analisis Algoritma C4.5 dengan PSO	20
3.6	Analisis Data	20
 BAB IV. HASIL DAN ANALISIS		
4.1	Hasil Penggunaan Algoritma C4.5	23
4.2	Hasil Penggunaan Algoritma C4.5 dengan Particle Swarm Optimization (PSO)	29
4.2.1	Inisialisasi Posisi pada Partikel	29
4.2.2	Bangkitkan Nilai Populasi dan Nilai <i>Velocity</i>	29
4.2.3	Nilai <i>Fitness</i> dan Nilai Kecepatan	30
4.3	Konfigurasi Hasil Percobaan	32
4.4	Implementasi	33
4.5	Pengukuran Performa Hasil Penelitian	33
4.5.1	Hasil Performa Nilai Confusion Matrix C4.5 dan C4.5 dengan PSO	35
4.5.2	Hasil Performa pada Kurva AUC	35
4.5	Analisis Peningkatan Akurasi	37
4.6	Analisis Performa Klasifikasi	38
 BAB V. KESIMPULAN DAN SARAN		
5.1	Kesimpulan	39
5.2	Saran	39
Daftar Pustaka		40

DAFTAR GAMBAR

Gambar 2.1	Pengelompokan TeknikKlasifikasi	10
Gambar 3.1	Model yang diusulkan	18
Gambar 3.2	Diagram Alir Agoritma C4.5 Menggunakan PSO	19
Gambar 4.1	Pohon Keputusan pada akar pertama	27
Gambar 4.2	Hasil Pohon Keputusan menggunakan <i>Rappid Miner</i>	28
Gambar 4.3	Implementasi Antarmuka pada saat mengambil data	33
Gambar 4.4	Implementasi Antarmuka menu <i>dashboard</i>	33
Gambar 4.5	Implementasi Antarmuka menu <i>analysis</i>	34
Gambar 4.6	Implementasi Antarmuka menu <i>classification</i> pada label gula darah	36
Gambar 4.7	Kurva Akurasi pada Konfigurasi Algoritma C4.5	37
Gambar 4.8	Grafik Akurasi pada Konfigurasi Algoritma C4.5 – PSO	37
Gambar 4.6	Grafik Perbandingan Performa <i>specificity</i> , <i>precision</i> , dan <i>recall</i> pada Konfigurasi Algoritma C4.5 dan C4.5-PSO	38

DAFTAR TABEL

Tabel 2.1	Faktor Resiko Diabetes Melitus	8
Tabel 2.2	Model <i>Confusion Matrix</i>	14
Tabel 3.1	Tabel Pasien Penderita Penyakit Gula Darah	21
Tabel 3.2	Kategori Atribut Data Penyakit Gula Darah	22
Tabel 4.1	Jumlah Kasus dari tiap Atribut	23
Tabel 4.2	Hasil Perhitungan Entropi dan Gain	26
Tabel 4.3	Tabel Atribut yang diinisialisasikan	29
Tabel 4.4	Inisialisasi nilai posisi partikel pada atribut yang digunakan	30
Tabel 4.5	Konfigurasi nilai parameter PSO	31
Tabel 4.6	Nilai <i>Fitness</i>	31
Tabel 4.7	Konfigurasi Percobaan	32
Tabel 4.8	Hasil Konfigurasi Percobaan	32
Tabel 4.9	Hasil Nilai <i>Confusion Matrix</i> pada Percobaan I	35
Tabel 4.10	Hasil Konfigurasi <i>Accuracy</i>	36

BAB I. PENDAHULUAN

1.1 Latar Belakang Masalah

Tahun 2025 Indonesia di prediksi menjadi 5 besar Negara dengan jumlah diabetes sebanyak 12,4 juta jiwa (Arisman., 2011). Jumlah dan jenis makanan yang merupakan bagian dari pola makan dianjurkan untuk sangat di atur oleh penderita Diabetes Mellitus, serta mengatur jadwal pemeriksaan kadar gula rutin ke dokter. Mengkonsumsi makanan tertentu dapat meningkatkan Kadar Gula Darah karena makanan yang dikonsumsi cenderung memiliki kandungan gula darah yang tidak terkontrol (Tandra., 2009).

Faktor yang mempengaruhi diagnosis yang tidak akurat pada gula darah, disebabkan oleh beberapa variabel yang sulit diukur seperti banyaknya faktor resiko yang terlalu banyak dan beberapa faktor resiko yang lainnya tidak relevan. Penggunaan jenis variabel yang banyak bukan berarti membuat kinerja menjadi lebih baik. Hal ini dapat menyebabkan kesalahan pengukuran dalam pemeriksaan kesehatan pasien dan hasil pemeriksaan yang tidak akurat. (Jinli He, 2020). Oleh karena itu, pada penelitian ini menggunakan data *historis* dari nilai gula darah sebagai data *input*.

Pekembangan teknologi dalam berbagai aspek sangat berpengaruh pada efisiensi pekerjaan masyarakat, tidak terkecuali di bidang kesehatan. Banyak instansi kesehatan rumah sakit, puskesmas, klinik yang mengatasi berbagai pasien Diabetes Mellitus. Kecanggihan teknologi yang berkembang saat ini memberikan manfaat dalam hal menyampaikan informasi yang cepat dan akurat khususnya di bidang kesehatan, baik untuk tenaga medis terlebih bagi penderita gula darah agar semakin mudah dalam memeriksa kesehatan. Jika data yang diberikan tidak memiliki informasi yang tepat dan tidak akurat akan menyebabkan kesalahan yang fatal seperti diagnosis yang salah, perencanaan layanan kesehatan yang kurang tepat, dan manajemen sistem kesehatan yang kacau (Stanislaw, 2016).

Perlu adanya efektifitas dalam memberikan informasi agar cepat dan akurat, salah satu teknologi informasi yang dapat dimanfaatkan dalam bidang kesehatan adalah data mining. Data mining merupakan istilah yang digunakan untuk

menguraikan penemuan pengetahuan di dalam database. Data Mining telah menjadi metodologi dasar untuk aplikasi komputasi di bidang medis. Data Mining dapat diterapkan dalam bidang kesehatan seperti mendiagnosis kanker payudara, penyakit jantung, Diabetes Mellitus dan lainnya. (Daniel L T, 2005)

Data mining merupakan kumpulan data yang besar bila diolah dengan benar akan memberikan informasi pengetahuan yang baru (Lakhsmi &Raghunandhan 2011). Prediksi dan klasifikasi data kerap digunakan dalam Data Mining untuk menggambarkan kelas data atau untuk memprediksi data di masa depan. Klasifikasi dalam memprediksi data telah banyak diterapkan agar suatu penyakit dapat di diagnosa dalam suatu penelitian. (Jain D, Singh V.,2018)

Sigit dalam penelitiannya melakukan prediksi pada penyakit Stroke (Sigit,2011). Penelitian lain pada penyakit *Diabetic Retinopathy* menggunakan metode Neural Network dan Particle Swarm Optimization menghasilkan akurasi 76.11%, sedangkan menggunakan Neural Network 71.76% (Toni.A, 2018). Penelitian yang pernah dilakukan oleh Muslim pada penyakit Diabetes Melitus, dengan membandingkan dua metode yaitu Algoritma C4.5 dan Algoritma C4.5 berbasis PSO menghasilkan tingkat akurasi 0,88% (Muslim,2018). Penelitian mengenai prediksi penyakit jantung yang dilakukan oleh Rusda, naik dengan akurasi sebesar 12,5% sebelumnya memiliki tingkat akurasi 81,25%, menjadi 93,75% (Rusda, 2014). Percobaan yang dilakukan oleh Warid dengan Algoritma klasifikasi yang berbeda yaitu Algoritma K-Nearest Neighbour menggunakan Particle Swarm Optimization (PSO) pada prediksi penyakit ginjal kronik, dalam penelitian ini PSO mampu meningkatkan nilai akurasi 18,5%, PSO terbukti lebih baik dalam memprediksi penyakit ginjal (Warid, 2018).

Pada teknik Particle Particle Swarm Optimization (PSO) memiliki kriteria dalam perhitungannya, dalam perhitungan PSO harus memiliki parameter seperti jumlah posisi, kecepatan, dan bobot inersia. Dalam perhitungan pengoptimasian dalam teknik PSO yaitu menghitung bobot pada atribut terhadap semua variabel yang digunakan dan dalam proses PSO juga dapat menseleksi jumlah atribut, dan menyeleksi fitur (Park, Lee & Choi, 2009). Pemberian bobot pada atribut yang digunakan algoritma C4.5 dengan menggunakan Particle Swarm Optimization ternyata dapat meningkatkan akurasi kinerja Algoritma C4.5 dan dapat pula

mempengaruhi pemilihan atribut yang tepat, apabila semakin banyak atribut yang digunakan maka akan memakan waktu yang lama dan memakan banyak biaya, hasilnya pun akan mengurangi tingkat akurasi dan kinerja akan menjadi lebih lambat (Andika, 2016).

Penelitian yang akan diangkat pada tesis ini akan membahas mengenai Pengoptimasian Bobot Atribut pada Algoritma C4.5 dengan menggunakan Particle swarm optimization dalam mendiagnosa pasien yang memiliki resiko Gula Darah Tinggi atau Normal dengan memilih kandidat kriteria pada penelitian sehingga diharapkan penelitian ini dapat memberikan kontribusi bagi tim medis khususnya di bidang kesehatan dengan tujuan agar dapat mengetahui dan dapat dilakukan penanganan lebih lanjut terhadap pasien yang memiliki Gula Darah Tinggi.

1.2 Rumusan Masalah

Berdasarkan Latar Belakang diatas, permasalahan dapat dirumuskan sebagai berikut :

1. Apakah Algoritma C4.5 menggunakan Particle swarm optimization dapat mengoptimasi bobot atribut pada prediksi Gula Darah?
2. Apakah dengan Particle Swarm Optimization, tingkat akurasi pada algoritma C4.5 dapat meningkat ?
3. Bagaimana perbandingan hasil akurasi ketika menggunakan Algoritma C4.5 dan Algoritma C4.5 menggunakan Particle swarm optimization?

1.3 Batasan Masalah

Batasan masalah yang akan dibahas pada penelitian ini adalah:

1. Dataset yang digunakan yaitu Pasien dengan Pengaruh Aktifitas Fisik Terhadap Kadar Gula Darah Diabetes Melitus.
2. Dataset yang digunakan memiliki dua class “Tinggi dan Normal”.
3. Fitur yang digunakan berjumlah 10 atribut yng akan dibagi menjadi 70% data training dan 30% data testing.
4. Mengukur dan menganalisis nilai parameter PSO (jumlah partikel) dalam peningkatan tingkat akurasi.

1.4 Tujuan Penelitian

Penelitian ini bertujuan untuk :

1. Optimasi Algoritma C4.5 dengan menggunakan *Particle Swarm Optimization* sehingga meningkatkan nilai akurasi kinerja dalam mendiagnosis dan memprediksi Gula Darah Tinggi atau Normal.
2. Membandingkan nilai akurasi pada algoritma C4.5 dan Algoritma C4.5 menggunakan *Particle Swarm Optimization*

1.5 Manfaat Penelitian

Manfaat penelitian ini adalah :

1. Manfaat dalam menggunakan PSO pada algoritma C4.5 agar tingkat akurasi kinerja pada hasil prediksi meningkat dan membantu para tenaga medis dalam memprediksi pasien penderita gula darah.
2. Dapat melihat perbedaan akurasi penggunaan algoritma C4.5 dan Algoritma C4.5 menggunakan PSO

3. Pada Bidang Kesehatan

Penelitian ini dapat memberikan kontribusi pada instansi kesehatan dalam mendiagnosis Gula Darah Tinggi atau Normal, sehingga dapat dilakukan tindakan lebih lanjut baik untuk pencegahan maupun pengobatan pada pasien yang terdeteksi.

4. Pada Pasien

Pasien akan mendapatkan manfaat deteksi lebih dini mengenai penyakit yang diderita agar mendapatkan pertolongan yang cepat dan tepat dari tim medis.

5. Pada Peneliti

Penelitian ini memberikan manfaat berupa pengimplementasian ilmu yang didapat baik dari dalam maupun diluar perkuliahan.

1.6 Sistematika Penulisan

Sistematika penulisan dibuat agar mempermudah penyusunan tesis dan garis besar uraian pada setiap bab. Sistematika penulisan disusun sebagai berikut:

1. BAB I - PENDAHULUAN

Bab I berisi tentang latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian dan sistematika penulisan.

2. BAB II - TINJAUAN PUSTAKA

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

3. BAB III – METODOLOGI PENELITIAN

Bab III 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 – ANALISA DAN PEMBAHASAN

Bab IV berisi tentang analisa dan pembahasan dari tiap-tiap blok diagram perencanaan rangkaian data-data hasil pengukuran.

5. BAB V – KESIMPULAN

Bab V berisi kesimpulan tentang hasil yang telah diperoleh serta merupakan jawaban dari tujuan yang ingin dicapai pada BAB I (Pendahuluan), akan tetapi masih bersifat sementara.

DAFTAR PUSTAKA

- A. Abraham, C.Grosan and V.Ramos, *Swarm Intelligence in Data mining*, Verlag Berlin Heidelberg: Springer,2006.
- Ana Moris,H.Peixoto, C.Coimbra, A.Abelha & J.Machado. (2017). Predicting the need of Neonatal Resuscitation using *Data mining. International Workshop on Healthcare Interoperability and Pervasive Intelligent Systems*. 571-576.
- Andika Dwi Hadiri. (2016). Penerapan Particle Swarm Optimization untuk Seleksi Atribut pada Metode Decision tree C4.5 untuk Persetujuan. KNIT-2 Nusa Mandiri
- Ari Puspita.(2016). Prediksi Kelahiran Bayi Secara Prematur dengan Mnegunakan Algoritma C4.5 Berbasis *Particle Swarm Optimization*. *Jurnal Teknik Informatika STMIK Antar Bangsa*.Vol. II.
- Bramer, Max. (2013f). *Principles Of Data mining*. London : Springer.pp: 176.
- Daniel T.Larose. (2005). *Discovering in Data mining, An Introduction to Data mining*. Willey Interscience
- Dinkes. (2015). *Profil Kesehatan Provinsi Sumatera Selatan*. Dinas Kesehatan Provinsi Sumatera Selatan.
- Dwi Ayu Nursela. (2010). Penerapan Algoritma C4.5 untuk Klasifikasi Tingkat Keganasan Kanker Payudara.
- Gorunescu, Florin. (2011). *Data mining : Concept and Techniques*. Verlag Berlin : Springer.
- Hameed S, Hassan R, Muhammad FF. Selection and classification of gene expression in autism disorder: Use of a combination of statistical filters and a GBPSO-SVM algorithm. *PloS one*. 2017; 12: e0187371.
- Hamid Abdul, Al-Ghamdi Abdullah, Naoman Amin, Madbouly Ayman, “A Comparative Analysis of Classification Algorithm for student college enrollment approval using *data mining*” pp. 1–8, 2012.
- Han, J., Pei J. & Kamber, M. (2011). *Data mining Concepts and Techniques*. Elsevier
- Haupt, R.L., Haupt, S.E. (2004). *Practical Genetic Algorithm*, Second Edition. Wiley
- Jain D, Singh V. (2018). Feature selection and classification systems for chronic diseaseprediction: A review.*Egyptian Informatics Journal*.

- Kemenkes. (2013). *Riset Kesehatan Dasar 2013*. Badan Penelitian dan Pengembangan Kesehatan, Kementerian Kesehatan RI, 65.
- Khun-Huang Chen, Khun-Jeng Wang, Kung-Min Wang, Melani-Andria Angelia. (2014). Applying particle swarm optimization-based decision tree classifier for cancer classification on gene expression data. *Applied Soft Computing*, 773-780.
- Lakshmi, B.N., Raghunandhan, G.H. (2011). A conceptual overview of *data mining*. *Proceedings of the National Conference on Innovations in Emerging Technology*, 27-32.
- Lis, Saumi., (2016). Penerapan Particle Swarm (PSO) untuk seleksi Atribut dalam meningkatkan akurasi prediksi Diagnosis Penyakit Hepatitis dengan Metode Algoritma C4.5. *SWABUMI* Vol. IV No.1.
- Mulkan, S. (2015). Penerapan *Particle Swarm Optimization* Untuk Menentukan Kredit Kepemilikan Rumah dengan Menggunakan Algoritma C4.5. *Jurnal Techno Nusa Mandiri* Vol. XII.
- M A Muslim et al. (2018) *Journal Physics.: Conference Series*. 983 012063
- Park, T. S., Lee, J. H., & Choi, B. (2009). Optimization for Artificial Neural Network with Adaptive Inertial Weight of *Particle Swarm Optimization* Cognitive Informatics, *IEEE International Conference*, 481- 485.
- Ronak Sumbaly, N. Ishnusri, S. Jeyalatha. (2014). Diagnosis of Breast Cancer using *Decision tree Data mining* Technique. *International Journal of Computer Application*, Vol.98 No.10, 16-24.
- Rusda Wajhillah, (2014). Optimasi Algoritma Klasifikasi C4.5 berbasis *Particle Swarm Optimization* untuk Prediksi Penyakit Jantung. *SWABUMI*, Vol.1.
- Mohammad Shafenoor Amin, Yin Kia Chiam, Kasturi Dewi Varathan (2018). *Telematics and Informatics*. Identification of significant features and *data mining* techniques in predicting heart disease.
- Stainlaw K., Dariusz M., Pawel K., Bozena MM., Daniel K. (2016). Beyond Databases, Architectures and Structures, *Advanced Technologies for Data Mining and Knowledge Discovery*. Springer. Hal.533.
- Sunjana, "Klasifikasi Data Nasabah Sebuah Asuransi Menggunakan Algoritma C4.5," Seminar Nasional Aplikasi Teknologi Informasi, pp. D31-D34, Juni 2010.
- T. H. Rahmadya and Prabowo P. W Herlawati (2013), Penerapan *Data mining* Dengan Matlab. Bandung : Rekayasa Sains: Andi.

Vercellis, Carlo. (2009). Business Intelligent : *Data mining* and Optimization for Decision Making. Shouther Gate, Gate, Chichester, West Sussex : John Willey & Sons, Ltd.

WHO. (2007). Pencegahan dan pengendalian infeksi Saluran Pernafasan Akut (ISPA) yang Cenderung Menjadi Epidemi dan Pandemi di Fasilitas Pelayan Kesehatan. Hal. 6.

Warid,Y (2018). Algoritma K-Nearest Neighbour Berbasis Particle Swarm Optimization untuk Prediksi Penyakit Ginjal Kronik. Jurnal Teknik Elektro CosPhi, Vol.2 No.2.

Witten, I. H., Frank, E., & Hall, M. A. (2011). *Data mining* Practical Machine Learning Tools and Techniques. United States: Elsevier Inc