

**PENINGKATAN AKURASI KLASIFIKASI PENYAKIT  
JANTUNG MENGGUNAKAN BACKPROPAGATION  
DENGAN SELEKSI FITUR ABC DAN K-NN**



**OLEH:**

**PANDITO DEWA PUTRA  
09042681721002**

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

**PENINGKATAN AKURASI KLASIFIKASI PENYAKIT  
JANTUNG MENGGUNAKAN BACKPROPAGATION  
DENGAN SELEKSI FITUR ABC DAN K-NN**

**TESIS**

**Diajukan Untuk Melengkapi Salah Satu Syarat  
Memperolah Gelar Magister**



**OLEH:**

**PANDITO DEWA PUTRA  
09042681721002**

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

## LEMBAR PENGESAHAN

# PENINGKATAN AKURASI KLASIFIKASI PENYAKIT JANTUNG MENGGUNAKAN BACKPROPAGATION DENGAN SELEKSI FITUR ABC DAN K-NN

## TESIS

Diajukan Untuk Melengkapi Salah Satu Syarat  
Memperoleh Gelar Magister

OLEH:  
**PANDITO DEWA PUTRA**  
**09042681721002**

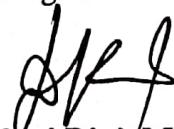
Pembimbing I



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

Palembang, 21 Januari 2021

Pembimbing II



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



## HALAMAN PERSETUJUAN

Pada hari Kamis tanggal 14 Januari 2021 telah dilaksanakan ujian sidang Tesis secara daring oleh Magister Ilmu Komputer Fakultas Ilmu Komputer Universitas Sriwijaya.

Nama : Pandito Dewa Putra

N I M : 09042681721002

Judul : Peningkatan Akurasi Klasifikasi Penyakit Jantung Menggunakan  
Backpropagation Dengan Seleksi Fitur ABC dan K-NN

### 1. Pembimbing I

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



### 2. Pembimbing II

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



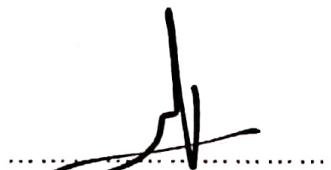
### 3. Penguji I

Dr. Ermatita, M.Kom.  
NIP. 196709132006042001



### 4. Penguji II

Dr. Abdiansah, S.Kom., M.Cs.  
NIP. 198410012009121005



Mengetahui,

Koordinator Program Studi Magister Ilmu Komputer



## LEMBAR PERNYATAAN

Yang bertanda tangan di bawah ini :

Nama : Pandito Dewa Putra  
NIM : 09042681721002  
Program Studi : Magister Ilmu Komputer  
Judul Tesis : Peningkatan Akurasi Klasifikasi Penyakit Jantung  
Menggunakan Backpropagation Dengan Seleksi Fitur  
ABC Dan K-NN

Hasil Pengecekan Software iThenticate/Turnitin : 8 %

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, 25 Januari 2021

Pandito Dewa Putra  
NIM. 09042681721002

## KATA PENGANTAR

Puji dan syukur penulis haturkan kehadirat Allah SWT yang telah melimpahkan rahmat, hidayah, dan kesehatan kepada penulis sehingga dapat menyelesaikan Tesis yang berjudul **“Peningkatan Akurasi Klasifikasi Penyakit Jantung Menggunakan Backpropagation Dengan Seleksi Fitur ABC dan K-NN”**.

Pada kesempatan ini, penulis juga hendak mengucapkan banyak terimakasih kepada pihak-pihak yang telah banyak membantu penyelesaian Tesis ini, diantaranya :

1. Orang tua, Bapak Mulhadi dan Ibu Nilem Sari.
2. Saudara kandung, Mbak Septiani Fransisca, S.E., M.Si., Ak.
3. Kakak ipar, Kak Bob Algadrie, S.E.
4. Dosen Pembimbing 1, Bapak Dr. Ir. Sukemi, M.T.
5. Dosen Pembimbing 2, Ibu Dian Palupi Rini, M.Kom., Ph.D.
6. Pejuang Tesis Indra Irawan dan Lia Andiani

Penulis menyadari bahwa masih banyak kekurangan dalam penulisan Tesis ini. Oleh karena itu, kritik dan saran yang membangun sangat diharapkan demi kemajuan karya tulis khususnya yang berkaitan dengan Tesis ini. Penulis berharap semoga Tesis ini dapat bermanfaat bagi semua pihak khususnya Magister Ilmu Komputer Universitas Sriwijaya.

Palembang, 25 Januari 2021

Penulis

# **PENINGKATAN AKURASI KLASIFIKASI PENYAKIT JANTUNG MENGGUNAKAN BACKPROPAGATION DENGAN SELEKSI FITUR ABC DAN K-NN**

## **ABSTRACT**

One of the non-communicable diseases that is susceptible to occur, especially when an individual is of a productive age, is heart disease. Death cases caused by heart disease based on data obtained from the World Health Organization (WHO) experienced more than 17 million victims in the world, a large enough number so that is the basis for researchers to process medical data on heart disease. The method to be applied for classification data mining was using the backpropagation algorithm because this algorithm had a simplicity and a fairly good performance. The backpropagation algorithm worked by adjusting the interconnected weights of neurons to achieve the minimum error between the output of the classification results and the real output. However, the Backpropagation Algorithm in achieving convergence tended to be slow when it got optimal accuracy, so the artificial bee colony algorithm was applied which usually converges to the global optimal efficiently to be used as feature selection from the applied Cleveland heart disease dataset. Then the k-nearest neighbor algorithm would be used to get the fitness value during the feature selection process because it was effective in reducing data dimensions while maintaining high classification accuracy. The results of the classification experiment with backpropagation without a combination resulted in an accuracy of 92.88%, while the combination classification experiment with backpropagation with artificial bee colony and k-nearest neighbor experienced an increase in accuracy by 95.93%.

Keywords: heart disease, Cleveland, backpropagation, artificial bee colony, k-NN

# **PENINGKATAN AKURASI KLASIFIKASI PENYAKIT JANTUNG MENGGUNAKAN BACKPROPAGATION DENGAN SELEKSI FITUR ABC DAN K-NN**

## **ABSTRAK**

Salah satu penyakit tidak menular yang rentan terjadi terutama saat seorang individu berada pada usia produktif yaitu penyakit jantung (*Heart Disease*). Kasus meninggal dunia yang diakibatkan penyakit jantung berdasarkan data yang diperoleh dari Organisasi Kesehatan Dunia (WHO) dialami lebih dari 17 juta korban di dunia, jumlah yang cukup besar sehingga hal ini lah yang mendasari peneliti untuk mengolah data medis penyakit jantung. Metode yang akan diterapkan untuk penambangan data klasifikasi yaitu menggunakan algoritma *backpropagation* sebab algoritma ini memiliki kesederhanaan dan kinerjanya yang cukup baik. Algoritma *backpropagation* mempunyai cara kerja dengan menyesuaikan bobot yang saling terhubung antara neuron untuk mencapai kesalahan minimum antara keluaran hasil klasifikasi dengan keluaran yang nyata. Namun *Algoritma Backpropagation* dalam mencapai konvergen cenderung lambat saat mendapat akurasi optimal, sehingga diterapkanlah algoritma *artificial bee colony* yang biasanya konvergen ke global optimal secara efisien untuk digunakan sebagai seleksi fitur dari dataset penyakit jantung (*cleveland heart disease*) yang diterapkan. Kemudian akan digunakan algoritma k-nearest neighbor untuk mendapatkan nilai *fitness* saat proses seleksi fitur karena efektif mereduksi dimensi data dengan tetap menjaga akurasi klasifikasi yang tinggi. Hasil percobaan klasifikasi dengan *backpropagation* tanpa kombinasi menghasilkan akurasi 92,88%, sedangkan percobaan klasifikasi kombinasi *backpropagation* dengan *artificial bee kolony* dan *k-nearest neighbor* mengalami peningkatan akurasi sebesar 95.93%.

Kata kunci : penyakit jantung, cleveland, *backpropagation*, ABC, k-NN.

## DAFTAR ISI

	Halaman
<b>LEMBAR PENGESAHAN</b>	<b>iii</b>
<b>HALAMAN PERSETUJUAN</b>	<b>iv</b>
<b>LEMBAR PERNYATAAN</b>	<b>v</b>
<b>KATA PENGANTAR</b>	<b>vi</b>
<b>DAFTAR ISI</b>	<b>ix</b>
<b>DAFTAR TABEL</b>	<b>xiii</b>
<b>DAFTAR GAMBAR</b>	<b>xv</b>
<b>DAFTAR LAMPIRAN</b>	<b>xvii</b>
<b>BAB I PENDAHULUAN</b>	<b>1</b>
1.1 Latar Belakang Masalah	1
1.2 Perumusan Masalah	4
1.3 Batasan Masalah	4
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	5
1.6 Metodologi Penelitian	5
<b>BAB II TINJAUAN PUSTAKA</b>	<b>7</b>
2.1 Tinjauan Penelitian	7
2.2 Atribut Heart Disease Dataset	12
2.3 Data Mining	12
2.4 Klasifikasi	12
2.5 Seleksi Fitur	13
2.6 Artificial Bee Colony	13
2.7 k-Nearest Neighbor (k-NN)	15
2.8 Backpropagation	16
2.9 Evaluasi Model	20
<b>BAB III METODOLOGI PENELITIAN</b>	<b>23</b>
3.1 Tahap Penelitian	23
3.2 Persiapan Data	25
3.3 Pemilihan Fitur menggunakan Artificial Bee Colony dan k-NN	26
3.4 Klasifikasi menggunakan Backpropagation	29

3.5 Validasi Data	30
3.6 Validasi Artificial Bee Colony dan k-NN	31
3.7 Validasi Backpropagation	31
3.8 Simulasi Sampel Fitur Seleksi ABC dan k-NN	32
3.9 Simulasi Klasifikasi dengan Backpropagation	38
3.10 Simulasi Coffusion Matrix	45
<b>BAB IV HASIL DAN ANALISIS</b>	<b>47</b>
4.1 Hasil Validasi Backpropagation	47
4.1.1 Hasil Validasi Backpropagation 5 hidden layer dan 100 epoch	47
4.1.2 Hasil Validasi Backpropagation 10 hidden layer dan 250 epoch	49
4.1.3 Hasil Validasi Backpropagation 15 hidden layer dan 500 epoch	51
4.1.4 Hasil Validasi Backpropagation 20 hidden layer dan 700 epoch	53
4.1.5 Hasil Validasi Backpropagation 25 hidden layer dan 1000 epoch	55
4.1.6 Hasil Analisis Validasi Backpropagation	57
4.2 Hasil Validasi Seleksi Fitur ABC dan k-NN dengan Backpropagation	58
4.2.1 Hasil Seleksi Fitur ABC dan k-NN dengan 5 iterasi	59
4.2.1.1 Hasil ABC-kNN dan Backpropagation dengan 5 hidden layer	60
4.2.1.2 Hasil ABC-kNN dan Backpropagation dengan 10 hidden layer	61
4.2.1.3 Hasil ABC-kNN dan Backpropagation dengan 15 hidden layer	63
4.2.1.4 Hasil ABC-kNN dan <i>Backpropagation</i> dengan 20 hidden layer	65
4.2.1.5 Hasil ABC-kNN dan <i>Backpropagation</i> dengan 25 hidden layer	67
4.2.1.6 Hasil Analisis BP dengan Seleksi Fitur ABC dan k-NN pada 5 iterasi	68
4.2.2 Hasil Seleksi Fitur ABC dan k-NN dengan 10 iterasi	70
4.2.2.1 Hasil ABC-kNN dan Backpropagation dengan 5 hidden layer	71
4.2.2.2 Hasil ABC-kNN dan Backpropagation dengan 10 hidden layer	72
4.2.2.3 Hasil ABC-kNN dan Backpropagation dengan 15 hidden layer	74
4.2.2.4 Hasil ABC-kNN dan <i>Backpropagation</i> dengan 20 hidden layer	75
4.2.2.5 Hasil ABC-kNN dan <i>Backpropagation</i> dengan 25 hidden layer	77
4.2.2.6 Hasil Analisis BP dengan Seleksi Fitur ABC dan k-NN pada 10 iterasi	78
4.2.3 Hasil Seleksi Fitur ABC dan k-NN dengan 30 iterasi	79

4.2.3.1 Hasil ABC-kNN dan Backpropagation dengan 5 hidden layer	80
4.2.3.2 Hasil ABC-kNN dan Backpropagation dengan 10 hidden layer	81
4.2.3.3 Hasil ABC-kNN dan Backpropagation dengan 15 hidden layer	83
4.2.3.4 Hasil ABC-kNN dan <i>Backpropagation</i> dengan 20 hidden layer	84
4.2.3.5 Hasil ABC-kNN dan <i>Backpropagation</i> dengan 25 hidden layer	85
4.2.3.6 Hasil Analisis BP dengan Seleksi Fitur ABC dan k-NN pada 30 iterasi	86
4.2.4 Hasil Seleksi Fitur ABC dan k-NN dengan 50 iterasi	88
4.2.4.1 Hasil ABC-kNN dan Backpropagation dengan 5 hidden layer	89
4.2.4.2 Hasil ABC-kNN dan Backpropagation dengan 10 hidden layer	90
4.2.4.3 Hasil ABC-kNN dan Backpropagation dengan 15 hidden layer	92
4.2.4.4 Hasil ABC-kNN dan <i>Backpropagation</i> dengan 20 hidden layer	94
4.2.4.5 Hasil ABC-kNN dan <i>Backpropagation</i> dengan 25 hidden layer	96
4.2.4.6 Hasil Analisis BP dengan Seleksi Fitur ABC dan k-NN pada 50 iterasi	97
4.2.5 Hasil Seleksi Fitur ABC dan k-NN dengan 100 iterasi	98
4.2.5.1 Hasil ABC-kNN dan Backpropagation dengan 5 hidden layer	99
4.2.5.2 Hasil ABC-kNN dan Backpropagation dengan 10 hidden layer	100
4.2.5.3 Hasil ABC-kNN dan Backpropagation dengan 15 hidden layer	102
4.2.5.4 Hasil ABC-kNN dan <i>Backpropagation</i> dengan 20 hidden layer	103
4.2.5.5 Hasil ABC-kNN dan <i>Backpropagation</i> dengan 25 hidden layer	105
4.2.5.6 Hasil Analisis BP dengan Seleksi Fitur ABC dan k-NN pada 100 iterasi	106
4.3 Hasil Perbandingan Klasifikasi BP dengan Seleksi Fitur ABC dan k-NN	108
4.4 Hasil Pengujian BP Berdasarkan Learning Rate dan Hidden Layer	110
4.4.1 Hasil Pengujian BP Berdasarkan <i>Learning Rate</i>	110
4.4.2 Hasil Pengujian BP Berdasarkan Hidden Layer	111
4.5 Hasil Pengujian BP dengan ABC-kNN Berdasarkan Learning Rate dan Hidden Layer	112

4.5.1 Hasil Pengujian BP dengan ABC-kNN Berdasarkan Learning Rate	112
4.5.2 Hasil Pengujian BP dengan ABC-kNN Berdasarkan Hidden Layer	113
<b>BAB V KESIMPULAN DAN SARAN</b>	<b>116</b>
5.1 Kesimpulan	116
5.2 Saran	116
<b>DAFTAR PUSTAKA</b>	<b>117</b>

## DAFTAR TABEL

	<b>Halaman</b>
<b>Tabel 2.1</b> Model Coffusion matrix	21
<b>Tabel 3.1</b> Cleveland Heart Disease	25
<b>Tabel 3.2</b> Struktur Seleksi Fitur	31
<b>Tabel 3.3</b> Struktur Klasifikasi Backpropagation	32
<b>Tabel 3.4</b> Fitur Seleksi ABC Inisiasi	33
<b>Tabel 3.5</b> Fitur Seleksi ABC Swap	33
<b>Tabel 3.6</b> Fitur Seleksi ABC Insert	34
<b>Tabel 3.7</b> Data training simulasi k-NN	35
<b>Tabel 3.8</b> Data Uji Simulasi k-NN	35
<b>Tabel 3.9</b> Hasil Euclidean Distance k-NN	36
<b>Tabel 3.10</b> Jarak Urutan Euclidean Distance	36
<b>Tabel 3.11</b> Kategori Dari Tetangga Terdekat	37
<b>Tabel 3.12</b> Hasil Simulasi Uji k-NN	37
<b>Tabel 3.13</b> Data sampel untuk simulasi Backpropagation	38
<b>Tabel 3.14</b> Bobot awal Random	38
<b>Tabel 3.15</b> Simulasi Coffusion Matrix	45
<b>Tabel 4.1</b> Coffusion Matrix dengan 5 hidden layer dan 100 epoch	48
<b>Tabel 4.2</b> hasil <i>backpropagation</i> dengan 5 hidden layer dan 100 epoch	48
<b>Tabel 4.3</b> Coffusion Matrix dengan 10 hidden layer dan 250 epoch	50
<b>Tabel 4.4</b> hasil <i>backpropagation</i> dengan 10 hidden layer dan 250 epoch	50
<b>Tabel 4.5</b> Coffusion Matrix dengan 15 hidden layer dan 500 epoch	52
<b>Tabel 4.6</b> hasil <i>backpropagation</i> dengan 15 hidden layer dan 500 epoch	52
<b>Tabel 4.7</b> Coffusion Matrix dengan 20 hidden layer dan 700 epoch	54
<b>Tabel 4.8</b> hasil <i>backpropagation</i> dengan 20 hidden layer dan 700 epoch	54
<b>Tabel 4.9</b> Coffusion Matrix dengan 25 hidden layer dan 1000 epoch	56
<b>Tabel 4.10</b> hasil <i>backpropagation</i> dengan 25 hidden layer dan 1000 epoch	56
<b>Tabel 4.11</b> Klasifikasi <i>Cleveland Heart Disease</i> dengan <i>Backpropagation</i>	57
<b>Tabel 4.12</b> hasil <i>backpropagation</i> dengan 5 hidden layer dan 100 epoch	60
<b>Tabel 4.13</b> hasil <i>backpropagation</i> dengan 10 hidden layer dan 250 epoch	62
<b>Tabel 4.14</b> hasil <i>backpropagation</i> dengan 15 hidden layer dan 500 epoch	63
<b>Tabel 4.15</b> hasil <i>backpropagation</i> dengan 20 hidden layer dan 700 epoch	65
<b>Tabel 4.16</b> hasil <i>backpropagation</i> dengan 25 hidden layer dan 1000 epoch	66
<b>Tabel 4.17</b> Klasifikasi BP dengan Seleksi Fitur ABC dan k-NN pada 5 iterasi	68
<b>Tabel 4.18</b> hasil <i>backpropagation</i> dengan 5 hidden layer dan 100 epoch	70
<b>Tabel 4.19</b> hasil <i>backpropagation</i> dengan 10 hidden layer dan 250 epoch	72

<b>Tabel 4.20</b> hasil <i>backpropagation</i> dengan 15 hidden layer dan 500 epoch	74
<b>Tabel 4.21</b> hasil <i>backpropagation</i> dengan 20 hidden layer dan 700 epoch	75
<b>Tabel 4.22</b> hasil <i>backpropagation</i> dengan 25 hidden layer dan 1000 epoch	76
<b>Tabel 4.23</b> Klasifikasi BP dengan Seleksi Fitur ABC dan k-NN pada 10 iterasi	78
<b>Tabel 4.24</b> hasil <i>backpropagation</i> dengan 5 hidden layer dan 100 epoch	80
<b>Tabel 4.25</b> hasil <i>backpropagation</i> dengan 10 hidden layer dan 250 epoch	82
<b>Tabel 4.26</b> hasil <i>backpropagation</i> dengan 15 hidden layer dan 500 epoch	83
<b>Tabel 4.27</b> hasil <i>backpropagation</i> dengan 20 hidden layer dan 700 epoch	84
<b>Tabel 4.28</b> hasil <i>backpropagation</i> dengan 25 hidden layer dan 1000 epoch	85
<b>Tabel 4.29</b> Klasifikasi BP dengan Seleksi Fitur ABC dan k-NN pada 30 iterasi	88
<b>Tabel 4.30</b> hasil <i>backpropagation</i> dengan 5 hidden layer dan 100 epoch	90
<b>Tabel 4.31</b> hasil <i>backpropagation</i> dengan 10 hidden layer dan 250 epoch	92
<b>Tabel 4.32</b> hasil <i>backpropagation</i> dengan 15 hidden layer dan 500 epoch	94
<b>Tabel 4.33</b> hasil <i>backpropagation</i> dengan 20 hidden layer dan 700 epoch	95
<b>Tabel 4.34</b> hasil <i>backpropagation</i> dengan 25 hidden layer dan 1000 epoch	96
<b>Tabel 4.35</b> Klasifikasi BP dengan Seleksi Fitur ABC dan k-NN pada 50 iterasi	98
<b>Tabel 4.36</b> hasil <i>backpropagation</i> dengan 5 hidden layer dan 100 epoch	100
<b>Tabel 4.37</b> hasil <i>backpropagation</i> dengan 10 hidden layer dan 250 epoch	102
<b>Tabel 4.38</b> hasil <i>backpropagation</i> dengan 15 hidden layer dan 500 epoch	103
<b>Tabel 4.39</b> hasil <i>backpropagation</i> dengan 20 hidden layer dan 700 epoch	104
<b>Tabel 4.40</b> hasil <i>backpropagation</i> dengan 25 hidden layer dan 1000 epoch	106
<b>Tabel 4.41</b> Klasifikasi BP dengan Seleksi Fitur ABC dan k-NN pada 100 iterasi	108
<b>Tabel 4.42</b> Pengujian BP Berdasarkan <i>Learning Rate</i>	111
<b>Tabel 4.43</b> Pengujian BP Berdasarkan <i>Hidden Layer</i>	111
<b>Tabel 4.44</b> Pengujian BP dengan ABC-kNN Berdasarkan <i>Learning Rate</i>	113
<b>Tabel 4.45</b> Pengujian BP dengan ABC-kNN Berdasarkan <i>Hidden Layer</i>	113
<b>Tabel 4.46</b> Penelitian tentang <i>Cleveland Heart Disease</i>	114

## DAFTAR GAMBAR

	<b>Halaman</b>
<b>Gambar 2.1</b> forward propagation dari signals	19
<b>Gambar 2.2</b> error propagation dalam backward direction	19
<b>Gambar 2.3</b> error propogasi menggunakan matriks bobot lain	19
<b>Gambar 3.1</b> Tahapan Penelitian Klasifikasi Penyakit Jantung menggunakan <i>Algoritma Backpropagation</i> dengan ABC dan k-NN	23
<b>Gambar 3.2</b> Flowchart Seleksi Fitur <i>Artificial Bee Colony</i> dan k-NN	27
<b>Gambar 3.3</b> Konsep seleksi fitur ABC dan k-NN	28
<b>Gambar 3.4</b> flowchart algoritma backpropagation	29
<b>Gambar 4.1</b> Grafik AUC dari <i>Backpropagation</i> dengan 5 hidden layer dan 100 epoch	49
<b>Gambar 4.2</b> Grafik AUC dari <i>Backpropagation</i> dengan 10 hidden layer dan 250 epoch	51
<b>Gambar 4.3</b> Grafik AUC dari <i>Backpropagation</i> dengan 15 hidden layer dan 500 epoch	53
<b>Gambar 4.4</b> Grafik AUC dari <i>Backpropagation</i> dengan 20 hidden layer dan 700 epoch	55
<b>Gambar 4.5</b> Grafik AUC dari <i>Backpropagation</i> dengan 25 hidden layer dan 1000 epoch	56
<b>Gambar 4.6</b> Klasifikasi <i>cleveland heart disease</i> dengan <i>Backpropagation</i>	57
<b>Gambar 4.7</b> Seleksi Fitur ABC dan k-NN dengan 5 iterasi	58
<b>Gambar 4.8</b> Grafik AUC dari <i>Backpropagation</i> dengan 5 hidden layer dan 100 epoch	59
<b>Gambar 4.9</b> Grafik AUC dari <i>Backpropagation</i> dengan 10 hidden layer dan 250 epoch	60
<b>Gambar 4.10</b> Grafik AUC dari <i>Backpropagation</i> dengan 15 hidden layer dan 500 epoch	62
<b>Gambar 4.11</b> Grafik AUC dari <i>Backpropagation</i> dengan 20 hidden layer dan 700 epoch	63
<b>Gambar 4.12</b> Grafik AUC dari <i>Backpropagation</i> dengan 25 hidden layer dan 1000 epoch	65
<b>Gambar 4.13</b> Klasifikasi BP dengan Seleksi Fitur ABC dan k-NN pada 5 iterasi	66
<b>Gambar 4.14</b> Seleksi Fitur ABC dan k-NN dengan 10 iterasi	67
<b>Gambar 4.15</b> Grafik AUC dari <i>Backpropagation</i> dengan 5 hidden layer dan 100 epoch	69
<b>Gambar 4.16</b> Grafik AUC dari <i>Backpropagation</i> dengan 10 hidden layer dan 250 epoch	70
<b>Gambar 4.17</b> Grafik AUC dari <i>Backpropagation</i> dengan 15 hidden layer dan 500 epoch	72

<b>Gambar 4.18</b> Grafik AUC dari <i>Backpropagation</i> dengan 20 hidden layer dan 700 epoch	73
<b>Gambar 4.19</b> Grafik AUC dari <i>Backpropagation</i> dengan 25 hidden layer dan 1000 epoch	75
<b>Gambar 4.20</b> Klasifikasi BP dengan Seleksi Fitur ABC dan k-NN pada 10 iterasi	76
<b>Gambar 4.21</b> Seleksi Fitur ABC dan k-NN dengan 30 iterasi	77
<b>Gambar 4.22</b> Grafik AUC dari <i>Backpropagation</i> dengan 5 hidden layer dan 100 epoch	79
<b>Gambar 4.23</b> Grafik AUC dari <i>Backpropagation</i> dengan 10 hidden layer dan 250 epoch	80
<b>Gambar 4.24</b> Grafik AUC dari <i>Backpropagation</i> dengan 15 hidden layer dan 500 epoch	82
<b>Gambar 4.25</b> Grafik AUC dari <i>Backpropagation</i> dengan 20 hidden layer dan 700 epoch	83
<b>Gambar 4.26</b> Grafik AUC dari <i>Backpropagation</i> dengan 25 hidden layer dan 1000 epoch	85
<b>Gambar 4.27</b> Klasifikasi BP dengan Seleksi Fitur ABC dan k-NN pada 30 iterasi	86
<b>Gambar 4.28</b> Seleksi Fitur ABC dan k-NN dengan 50 iterasi	87
<b>Gambar 4.29</b> Grafik AUC dari <i>Backpropagation</i> dengan 5 hidden layer dan 100 epoch	89
<b>Gambar 4.30</b> Grafik AUC dari <i>Backpropagation</i> dengan 10 hidden layer dan 250 epoch	90
<b>Gambar 4.31</b> Grafik AUC dari <i>Backpropagation</i> dengan 15 hidden layer dan 500 epoch	92
<b>Gambar 4.32</b> Grafik AUC dari <i>Backpropagation</i> dengan 20 hidden layer dan 700 epoch	93
<b>Gambar 4.33</b> Grafik AUC dari <i>Backpropagation</i> dengan 25 hidden layer dan 1000 epoch	95
<b>Gambar 4.34</b> Klasifikasi BP dengan Seleksi Fitur ABC dan k-NN pada 50 iterasi	96
<b>Gambar 4.35</b> Seleksi Fitur ABC dan k-NN dengan 100 iterasi	97
<b>Gambar 4.36</b> Grafik AUC dari <i>Backpropagation</i> dengan 5 hidden layer dan 100 epoch	99
<b>Gambar 4.37</b> Grafik AUC dari <i>Backpropagation</i> dengan 10 hidden layer dan 250 epoch	100
<b>Gambar 4.38</b> Grafik AUC dari <i>Backpropagation</i> dengan 15 hidden layer dan 500 epoch	102
<b>Gambar 4.39</b> Grafik AUC dari <i>Backpropagation</i> dengan 20 hidden layer dan 700 epoch	103
<b>Gambar 4.40</b> Grafik AUC dari <i>Backpropagation</i> dengan 25 hidden layer dan 1000 epoch	105
<b>Gambar 4.41</b> Klasifikasi BP dengan Seleksi Fitur ABC dan k-NN Pada 100 iterasi	108
<b>Gambar 4.42</b> Peningkatan Akurasi dengan Seleksi Fitur	109

## **DAFTAR LAMPIRAN**

**LAMPIRAN 1.** Source Code Python

**LAMPIRAN 2.** Publikasi Ilmiah

**LAMPIRAN 3.** Hasil Pengecekan Software Ithenticate/Turniting

**LAMPIRAN 4.** Form Konsultasi

**LAMPIRAN 5.** Form Revisi Ujian Tesis II

## **BAB I**

### **PENDAHULUAN**

Bab ini menjabarkan latar belakang penelitian ini dengan judul: “Seleksi Fitur ABC dan k-NN untuk Peningkatan Akurasi Klasifikasi dengan *Backpropagation* pada Penyakit Jantung”. Permasalahan Penyakit jantung dipilih menjadi topik dikarenakan penyakit jantung mengalami peningkatan setiap tahunnya. Rasio penderita penyakit jantung mencapai 15 dari 1000 orang atau sekitar lebih dari dua juta individu mengalami penyakit ini di Indonesia (<http://www.inaheart.org>, 2019).

#### **1.1 Latar Belakang Masalah**

Penyakit jantung (*heart disease*) merupakan penyakit tidak menular yang seringkali diderita individu pada usia produktif. Tanda-tanda munculnya penyakit jantung sulit dideteksi oleh penderitanya akibat kurangnya pengetahuan. Hal ini mengakibatkan faktor kematian menjadi kian tinggi (Sabransyah, M., dkk. 2017). Jantung sangat penting bagi tubuh, kehidupan itu sendiri tergantung pada kinerja jantung yang efisien. Penyakit jantung adalah penyakit yang didasarkan pada kinerja jantung. Beberapa faktor yang meningkatkan risiko penyakit jantung seperti: kurangnya aktivitas berolahraga, kolesterol, obesitas, merokok dan tekanan darah tinggi (Al-Milli, 2013).

Kasus meninggal dunia yang diakibatkan penyakit jantung berdasarkan data yang diperoleh dari Organisasi Kesehatan Dunia (WHO) dialami lebih dari 17 juta korban di dunia. Sementara itu Riset Kesehatan Dasar (Riskesdas) menunjukkan data penyakit jantung yang terus menanjak pada tahun 2018. Setidaknya terdapat 0.015% penderita penyakit ini di Indonesia dan terus meingkat dari waktu ke waktu (Indonesian Heart Association, 2019).

*Prognosis* atau prediksi suatu penyakit dapat dipengaruhi berbagai faktor atau gejala sebab hal itu merupakan suatu permasalahan yang kompleks dan rumit, bahkan dapat menyebabkan asumsi yang salah. Oleh karena itu, dilakukan suatu upaya agar menjadi penghubung antara pengetahuan dan pengalaman para

ahli dengan teknologi yang ada saat ini, sehingga membangun sistem yang cukup untuk mendukung proses diagnosis suatu penyakit (M Durairaj and V Revathi, 2015).

Bidang teknologi informasi saat ini menjadi sangat popular dan sering diterapkan diberbagai bidang, salah satunya yaitu teknik *data mining* yang dikombinasikan dengan bidang ilmu kesehatan. *Data mining* merupakan suatu bidang ilmu yang berkaitan dengan perhitungan secara komputasional terhadap berbagai macam data yang luas. *Data mining* dapat mengekstraksi pengetahuan yang bermanfaat dari kumpulan data-data dan memberikan hasil pengambilan keputusan yang tepat, salah satunya yaitu menggunakan data – data medis seperti data penyakit jantung. Penambangan data terhadap bidang kesehatan memiliki potensi besar untuk menemukan pola tersembunyi di antara kumpulan data tentang domain medis sehingga dapat berguna untuk diagnosis atau pengobatan penyakit (Kodati, 2018).

Teknik penambangan data yang populer saat ini, telah diterapkan pada pengklasifikasian data – data. Metode klasifikasi penambangan data yang sering digunakan saat ini yaitu *backpropagation neural network*. Algoritma *backpropagation* mengadopsi prinsip jaringan *multilayer feed-forward* yang memodelkan suatu fungsi dengan melakukan perhitungan kesalahan dari masing-masing neuron, kemudian memperbarui bobotnya setelah memproses grup data hingga nilai kesalahan berada dalam ambang batas (Desai, 2018). *Backpropagation* dikenal juga sebagai algoritma backpropagation error berdasarkan dari aturan pengetahuan perbaikan error (Kodati, 2018). Namun *Algoritma Backpropagation* dalam mencapai konvergen cenderung lambat saat mendapat akurasi optimal, serta jatuh ke minimum lokal dengan mudah (Jin Feihu and Shu Guang, 2012).

Teknik data mining sangat memerlukan hasil akurasi yang tinggi, untuk meningkatkan akurasi tersebut maka diterapkanlah *fitur selection* agar suatu algoritma mengalami peningkatan akurasi. Pada penelitian (Rahman, 2017) dan (Rao Haidi, 2019) menunjukkan bahwa dengan menerapkan *fitur selection* meningkatkan akurasi dari metode-metode klasifikasi yang diterapkan. Salah satu

metode pemilihan fitur yang sering digunakan yaitu dengan algoritma metaheuristik, yang merupakan algoritma untuk melakukan optimisasi dan juga terkadang dikombinasikan dengan algoritma lainnya untuk meningkatkan suatu akurasi yang dicapai. Salah satu metode metaheuristik yaitu *artificial bee Colony* dikombinasikan dengan algoritma *gradient boosting decision tree* dibandingkan dengan beberapa algoritma lain dan kemudian diuji dengan beberapa *dataset* (Rao Haidi, 2019). Pada dataset Haberman mencapai akurasi tertinggi (ABC-DT) 85.52%, kemudian pada *dataset* WDBC kembali mencapai akurasi tertinggi (ABC-DT) 97.18%. Selanjutnya pada *dataset* glass mencapai akurasi tertinggi (ABC-DT) 70.37%.

*Algoritma metaheuristic (harmony search (HS) berbasis enhanced, Genetic algorithm (GA), Particle swarm optimization (PSO))* dibandingkan dengan beberapa *dataset* (Zainuddin, 2016). Pada *Heart Dataset* akurasi tertinggi oleh GA 93.67%, kemudian pada Hepatitis *dataset* akurasi tertinggi oleh Enhanced HS 97.50%. pada *dataset* iris akurasi tertinggi oleh PSO 97.33%. Kemudian Enhanced HS unggul di *dataset* glass, ionosphere, vehicle, wine dan zoo. Selanjutnya *Genetic algorithm* unggul di *dataset* votes, kemudian untuk PSO unggul didataset pima.

*Algoritma artificial bee Colony* merupakan metode optimasi yang sangat populer, digunakan untuk proses pemilihan fitur dalam penelitian ini (Uzer Mustafa Serter, 2013). Database yang mengalami pemilihan fitur diklasifikasikan menggunakan SVM. Akurasi dengan hepatitis dataset mencapai 94.92%. Akurasi dengan *liver disorders dataset* mencapai 82.55%.

*Artificial Bee Colony* (ABC) yaitu algoritma yang seringkali dikonvergen ke global optimal secara efisien dan menghabiskan lebih sedikit waktu untuk pemilihan fitur (Rao Haidi, 2019). ABC memiliki keunggulan yaitu struktur yang lebih sederhana, parameter kontrol lebih sedikit dan mudah direalisasikan. Dalam mendapatkan nilai *fitness* pada *algoritma* ABC maka akan digunakan *algoritma k-Nearest Neighbor (k-NN)*. Kedua kerangka kerja *Algoritma* ABC maupun k-NN telah dimodifikasi dan digabungkan untuk membuat metode reduksi dimensi yang efektif, Hasil analisis ekspresi gen menunjukkan bahwa metode ABC - k-NN

efektif mereduksi dimensi data dengan tetap menjaga akurasi klasifikasi yang tinggi. (Prasartvit, et al. 2012).

Oleh sebab itu, berdasarkan penelitian sebelumnya, peneliti akan melakukan pengujian *algoritma backpropagation* dengan menggunakan objek data penyakit jantung (*cleveland heart disease*). Kemudian akan dilakukan proses seleksi fitur dengan *algoritma artificial bee colony* dan k-NN untuk meningkatkan akurasi pada analisis ini. Selanjutnya dibuat perbandingan menggunakan *algoritma backpropagation* itu sendiri tanpa seleksi fitur.

## 1.2 Perumusan Masalah

Penelitian ini akan membahas beberapa isu penting sesuai dengan penjelasan latar belakang di atas, sehingga dirumuskan masalah sebagai berikut:

1. Bagaimana mengatasi *algoritma backpropagation* yang cenderung lambat untuk mencapai konvergen.
2. Bagaimana seleksi fitur pada *algoritma backpropagation* dengan menggunakan *algoritma artificial bee colony* dan k-NN.
3. Bagaimana peningkatan kinerja dari *algoritma backpropagation* dibandingkan dengan algoritma kombinasi *backpropagation* dengan *artificial bee colony* dan k-NN.

## 1.3 Batasan Masalah

Klasifikasi penyakit jantung dengan memakai *algoritma backpropagation* merupakan batasan dalam penelitian, kemudian seleksi attribut menggunakan *algoritma artificial bee colony* dan k-NN. Dataset yang berasal dari UCI *machine learning repository* digunakan sebagai dataset dalam penelitian ini, lebih spesifiknya yaitu *cleveland heart disease dataset*.

## 1.4 Tujuan Penelitian

Penelitian ini diharapkan bertujuan sebagai berikut:

1. Peningkatan kinerja dari *algoritma backpropagation* dalam mencapai nilai konvergen yang maksimum.
2. Menggunakan algoritma optimasi *artificial bee colony* dan k-NN untuk seleksi atribut / *feature* dan *algoritma backpropagation* untuk melakukan klasifikasi data.
3. Membandingkan akurasi dari klasifikasi penyakit jantung menggunakan *algoritma backpropagation*, dan klasifikasi penyakit jantung menggunakan *algoritma backpropagation* dengan seleksi atribut berbasis *Artificial Bee Colony* dan k-NN.

## 1.5 Manfaat Penelitian

Sedangkan penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Memberi nilai tambah (*value added*) bagi penelitian di bidang data mining klasifikasi.
2. Memberikan alternatif bagi bidang kesehatan untuk melakukan klasifikasi penyakit jantung dengan menggunakan *algoritma backpropagation*.
3. Penelitian ini dapat dijadikan bahan referensi dalam peningkatan performa akurasi menggunakan sistem dengan metode kombinasi *backpropagation* dengan *artificial bee colony* dan k-NN.

## 1.6 Metodologi Penulisan

Sistematika penelitian dirancang agar penelitian memberikan gambaran yang lebih terarah dalam penelitian ini, diantaranya:

### 1. BAB I Pendahuluan

Latar belakang mengawali isi dari bab ini, kemudian dilanjutkan dengan perumusan masalah dan pemberian

batasan masalah. Sub bab berikutnya tujuan penelitian, dilanjutkan dengan manfaat penelitian.

## **2. BAB II**

### **Tinjauan Pustaka**

Penjelasan bab ini mengenai penelitian sebelumnya yang terkait dengan klasifikasi penyakit jantung menggunakan metode *backpropagation* dan *artificial bee colony* yang mengacu pada beberapa penelitian publikasi. Kemudian menjelaskan tentang penggunaan dataset yaitu cleveland heart disease. Tak lupa penulis menambahkan teori-teori yang sesuai dengan penelitian ini.

## **3. BAB III**

### **Metodologi Penelitian**

Tahapan terperinci serta langkah dalam pengumpulan dan penganalisaan klasifikasi penyakit jantung mengawali penjelasan dalam bab ini. Penjelasan *framework* algoritma *backpropagation* dijelaskan dalam metodologi penelitian ini. Penelitian ini juga menjelaskan model yang dipakai agar tercapai tujuan penelitian. Kemudian menjelaskan konsep dari fitur selection yang akan diterapkan dalam penelitian ini.

## **4. BAB IV**

### **Analisa dan Pembahasan**

Hasil testing uji yang dilakukan, akan dijelaskan dalam bab ini. Pengujian data kemudian ditelaah dengan berbagai rangkaian macam jenis validasi hasil maupun teknik tertentu.

## **5. BAB V**

### **Kesimpulan**

Hasil akhir yang kemudian disimpulkan akan dituangkan dalam bab ini, dan menjadi jawaban terhadap berbagai tujuan yang sebelumnya ingin dicapai dalam penelitian.

## DAFTAR PUSTAKA

- Al-Milli, Nabeel. 2013. *Backpropogation Neural Network For Prediction of Heart Disease*. Journal of Theoretical and Applied Information Technology, Vol.56 No.1 p.131-135. E-ISSN: 1817-3195.
- Amin, Mohammad Shafenoor., et al. 2019. *Identification Of Significant Features And Data Mining Techniques In Predicting Heart Disease*. Telematics and Informatics. Vol.36.
- Anggraeny, Fetty Tri dan Purbasari, Intan Yuniar. 2019. Jaringan Saraf Tiruan dan Modifikasinya Menggunakan Supervised Learning. ISBN: 978-623-7137-70-2.
- Arabasadi, Zeinab., et al. 2017. *Computer Aided Decision Making For Heart Disease Detection Using Hybrid Neural Network-Genetic Algorithm*. Computer Methods and Programs in Biomedicine. Vol.141.
- Chen, Xu., et al. 2020. *Fireworks Explosion Based Artificial Bee Colony For Numerical Optimization*. Vol.188.
- Desai, Shrinivas D., et al. 2018. *Back-Propagation Neural Network Versus Logistic Regression in Heart Disease Classification*. Advances in Intelligent Systems and Computing, Vol.702 ISBN: 978-981-13-0679-2.
- Feshki, Majid Ghonji and Shijani, Omid Sojoodi. 2016. *Improving the Heart Disease Diagnosis by Evolutionary Algorithm of PSO and Feed Forward Neural Network*. Artificial Intelligence and Robotics (IRANOPEN). ISBN: 978-1-5090-2169-7.
- Gokulnath, Chandra Babu and Shantharajah, S. P. 2018. *An Optimized Feature Selection Based On Genetic Approach And Support Vector Machine For Heart Disease*. Cluster Computing. Vol.22.
- Gorunescu, F (2011): *Data Mining Concept Model and Techniques*, BerlinSpringer, 14, 1 - 353.
- Han, Nan., et al. 2019. *A Novel Chinese Herbal Medicine Clustering Algorithm Via Artificial Bee Colony Optimization*. Vol.101.
- Indonesian Heart Association. 2019. Info Datin: Hari Jantung Sedunia. [http://www.inaheart.org/news\\_and\\_events/news/2019/9/26/press\\_release\\_world\\_heart\\_day\\_perki\\_2019](http://www.inaheart.org/news_and_events/news/2019/9/26/press_release_world_heart_day_perki_2019).

- Jiang, Ping and Chen, Jiejie. 2016. *Displacement Prediction Of Landslide Based On Generalized Regression Neural Networks With K-Fold Cross-Validation*. Vol.198.
- Jin, Feihu and Shu, Guang. 2012. *Back Propagation Neural Network Based On Artificial Bee Colony Algorithm*. International Forum on Strategic Technology (IFOST), ISBN: 978-1-4673-1772-6.
- Karayilan, Tulay dan Kilic, Ozkan. 2017. *Prediction of Heart Disease Using Neural Network*. 2017 International Conference on Computer Science and Engineering (UBMK), E-ISBN: 978-1-5386-0930-9.
- Khiarak, Jalil Nourmohammadi., et al. 2019. *New Hybrid Method For Heart Disease Diagnosis Utilizing Optimization Algorithm In Feature Selection*. Health and Technology.
- Kodati, Sarangam and Vivekanandam. 2018. *Analysis of Heart Disease using in Data Mining Tools Orange and Weka*. Global Journal of Computer Science and Technology, Vol.18 E-ISSN: 0975-4172. ISSN: 0975-4350.
- Latha, C. Beulah Christalin and Jeeva, S. Carolin. 2019. *Improving The Accuracy Of Prediction Of Heart Disease Risk Based On Ensemble Classification Techniques*. Informatics in Medicine Unlocked. Vol.16.
- Leema, N., at al. 2016. *Neural Network Classifier Optimization Using Differential Evolution With Global Information And Back Propagation Algorithm For Clinical Datasets*. Applied Soft Computing. Vol.49.
- M, Durairaj and V, Revathi. 2015. *Prediction Of Heart Disease Using Back Propagation MLP Algorithm*. International Journal Of Scientific & Technology Research, Vol.4 ISSN: 2277-8616.
- Mandrekar, Jayawant N. 2010. *Receiver Operating Characteristic Curve in Diagnostic Test Assessment*. Journal of Thoracic Oncology. Vol.5 ISSN : 1556-0864.
- Mienye, Ibomoiye Domor., et al. 2020. *Improved Sparse Autoencoder Based Artificial Neural Network Approach For Prediction Of Heart Disease*. Informatics in Medicine Unlocked. Vol.18.
- Mohan Senthilkumar., at al. 2019. *Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques*. IEEE Access, Vol.7 E-ISSN:2169-3536.

- Negrov, D., et al. 2017. *An Approximate Backpropagation Learning Rule For Memristor Based Neural Networks Using Synaptic Plasticity*. Neurocomputing. Vol.237.
- Olaniyi, Ebenezer Obaloluwa., at al. 2015. *Heart Diseases Diagnosis Using Neural Networks Arbitration*. International Journal of Intelligent Systems Technologies and Applications, Vol.12.
- Orkcu, H Hasan and Bal, Hasan. 2011. *Comparing performances of backpropagation and genetic algorithms in the data classification*. Expert Systems with Applications, Vol.38.
- Rahman, L., Setiawan, N.A. and Permanasari, A.E., 2017. *Feature Selection Methods in Improving Accuracy of Classifying Students' Academic Performance*. 2nd International Conferences on Information Technology, Information Systems and Electrical Engineering (ICITISEE). ISBN: 978-1-5386-0657-5.
- Rao, Haidi., et al. 2019. *Feature Selection Based On Artificial Bee Colony And Gradient Boosting Decision Tree*. Applied Soft Computing. Vol.74.
- Reddy, G. Thippa., et al. 2019. *Hybrid Genetic Algorithm And A Fuzzy Logic Classifier For Heart Disease Diagnosis*. Evolutionary Intelligence.
- Sabransyah, M., dkk. 2017. Aplikasi Metode Naive Bayes dalam Prediksi Risiko Penyakit Jantung. Jurnal EKSPONENSIAL, Vol.8 ISSN: 2085-7829.
- Sarangi, Partha Pratim., et al. 2014. *Training a Feed-Forward Neural Network Using Artificial Bee Colony with Back-Propagation Algorithm*. Advances in Intelligent Systems and Computing, vol 243. ISBN: 978-81-322-1664-3.
- Sharma, Purushottam and Saxena, Kanak. 2017. *Application Of Fuzzy Logic And Genetic Algorithm In Heart Disease Risk Level Prediction*. International Journal of System Assurance Engineering and Management. Vol.8.
- Srivastava, Shweta. 2014. *Weka: A Tool for Data preprocessing, Classification, Ensemble, Clustering and Association Rule Mining*. International Journal of Computer Applications. Vol.88.
- Sukemi., at al. 2012. *Priority Based Computation A Study on Paradigm Shift on Real Time Computation*. IEEE International Conference on Computational Intelligence and Cybernetics (CyberneticsCom), Vol.12 ISBN: 978-1-4673-0892-2.

- Uzer, Mustafa Serter., et al. 2013. *Feature Selection Method Based on Artificial Bee Colony Algorithm and Support Vector Machines for Medical Datasets Classification*. Vol.2013.
- Yoo, Illhoi., et al. 2011. *Data Mining in Healthcare and Biomedicine: A Survey of the Literature*. Journal of Medical Systems. Vol.36.
- Zainuddin, Zarita., et al. 2016. *An Enhanced Harmony Search Based Algorithm For Feature Selection: Applications In Epileptic Seizure Detection And Prediction*. Computers & Electrical Engineering. Vol.53.
- Zamorano, Francisco Ortega., et al. 2017. *FPGA Implementation of Neurocomputational Models: Comparison Between Standard Back-Propagation and C-Mantec Constructive Algorithm*. Vol.46.
- Zubcoff, Jose dan Trujillo, Juan. 2006. *Conceptual Modeling for Classification Mining in Data Warehouses*. Data Warehousing and Knowledge Discovery, E-ISBN: 978-3-540-37737-5.
- Aci, Mehmet., et al. 2010. *A Hybrid Classification Method Of K Nearest Neighbor, Bayesian Methods And Genetic Algorithm*. Expert Systems with Applications. Vol.37.
- Yadav, Sanjay dan Shukla, Sanyam. 2016. *Analysis of k-Fold Cross-Validation over Hold-Out Validation on Colossal Datasets for Quality Classification*. Proc. - 6th Int. Adv. Comput. Conf. IACC, no. Cv, pp. 78–83, doi: 10.1109/IACC.2016.25
- Prasarvit, Thananan., et al. 2012. *Reducing bioinformatics data dimension with ABC-kNN*. Neurocomputing, vol. 116, pp. 367–381, 2013, doi: 10.1016/j.neucom.2012.01.045.
- Iftikhar, Saman., et al. 2017. An Evolution Based Hybrid Approach For Heart Diseases Classification And Associated Risk Factors Identification. Vol.28, PP.3451-3455.
- Pouriyeh, Seyedamin., et al. 2017. A Comprehensive Investigation and Comparison of Machine Learning Techniques in the Domain of Heart Disease. 22nd IEEE Symposium on Computers and Communication (ISCC 2017). Vol.1, PP. 204-207.
- Mary, Stella dan Sebastian, Shoney. 2019. Predicting Heart Ailment In Patients With Varying Number Of Features Using Data Mining Techniques. International Journal of Electrical and Computer Engineering (IJECE). Vol.9. PP.2675-2681.

- Verma, Luxmi., et al. 2016. A Hybrid Data Mining Model to Predict Coronary Artery Disease Cases Using Non-Invasive Clinical Data. *Journal of Medical Systems*. Vol.178.
- Gupta, Ankur., et al. 2019. MIFH: A Machine Intelligence Framework for Heart Disease Diagnosis. vol. 8, pp. 14659-14674.
- Al-Yarimi, F.A.M., et al. 2020. Feature Optimization By Discrete Weights For Heart Disease Prediction Using Supervised Learning. *Soft Comput* (2020).
- Mienye, D. I., et al. 2020. An Improved Ensemble Learning Approach For The Prediction Of Heart Disease Risk. *Informatics in Medicine Unlocked*. Vol.20.
- Latha, C. B. C dan Jeeva, Carolin. 2019. Improving The Accuracy Of Prediction Of Heart Disease Risk Based On Ensemble Classification Techniques. *Informatics in Medicine Unlocked*. Vol.16.
- Javeed, Ashir., et al. 2019. An Intelligent Learning System Based on Random Search Algorithm and Optimized Random Forest Model for Improved Heart Disease Detection. *IEEE Access*, vol. 7, pp. 180235-180243.
- Khennou, Fadoua., et al. 2019. A Machine Learning Approach: Using Predictive Analytics to Identify and Analyze High Risks Patients with Heart Disease. *International Journal of Machine Learning and Computing*. Vol. 9. No. 6.
- Maashi, S. Mashael. 2020. Analysis Heart Disease Using Machine Learning. *Multi-Knowledge Electronic Comprehensive Journal For Education And Science Publications (MECSJ)*. ISSN: 2616-9185.
- Moturi, Sireesha., et al. 2020. Classification Model for Prediction of Heart Disease using Correlation Coefficient Technique. *International Journal of Advanced Trends in Computer Science and Engineering*. Vol.9. No.2.
- Javid, Irfan., et al. 2020. Enhanced Accuracy of Heart Disease Prediction using Machine Learning and Recurrent Neural Networks Ensemble Majority Voting Method. (IJACSA) *International Journal of Advanced Computer Science and Applications*. Vol.11. No.3.
- Yudianto, M. D., et al. 2020. A Feature-Driven Decision Support System for Heart Disease Prediction Based on Fisher's Discriminant Ratio and Backpropagation Algorithm. Vol.11. E-ISSN: 2541-5832.
- Krstinić, Damir., et al. 2020. Multi-Label Classifier Performance Evaluation With Confusion Matrix. *Computer Science & Information Technology (CS & IT)*. pp. 01-14.