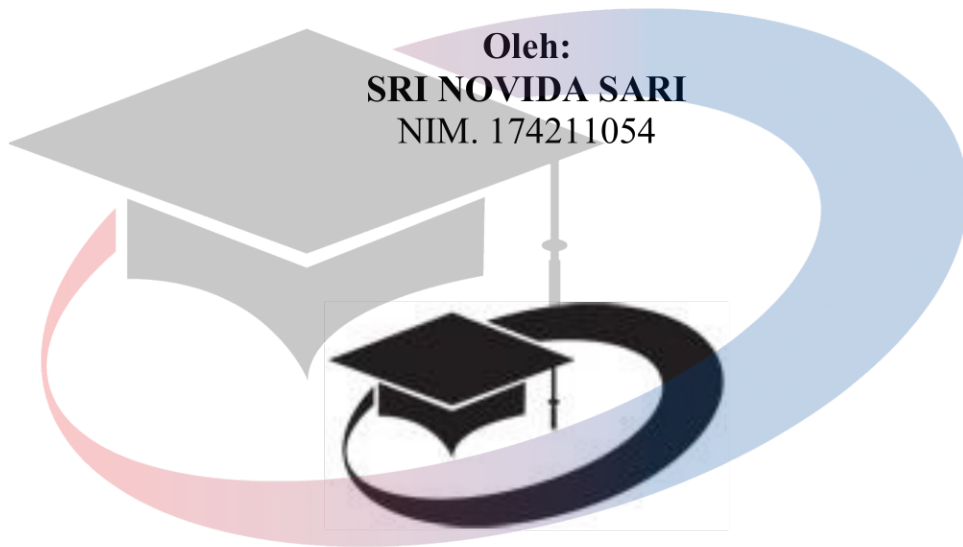


**OPTIMASI *GENETIC ALGORITHM* MELALUI
PENANGANAN INCOMPLETE DATA DAN REDUKSI**

TESIS



Oleh:
SRI NOVIDA SARI
NIM. 174211054

**UNIVERSITAS
MIKROSKIL**

**PROGRAM STUDI MAGISTER TEKNOLOGI INFORMASI
SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN**

KOMPUTER

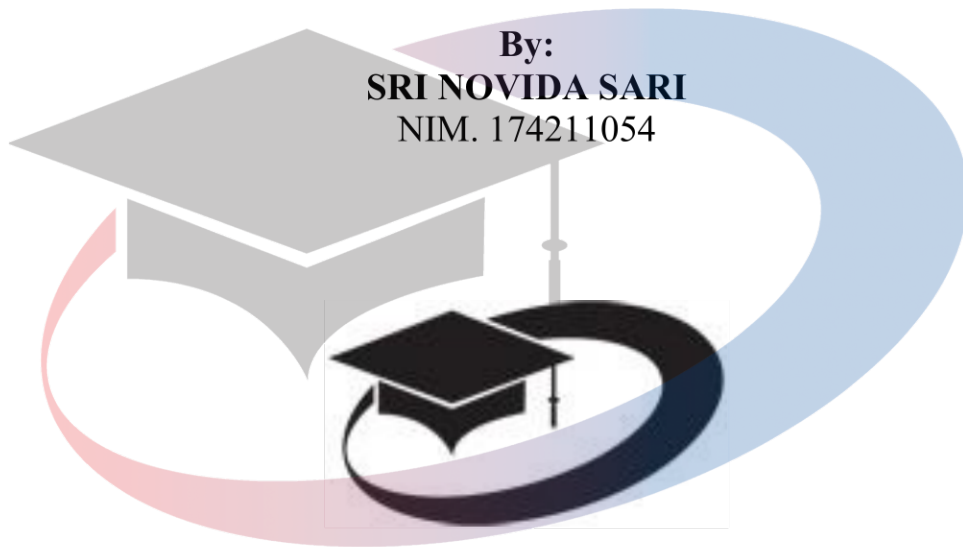
MIKROSKIL

MEDAN

2020

**GENETIC ALGORITHM OPTIMIZATION THROUGH
HANDLING OF INCOMPLETE DATA AND REDUCTION**

THESIS



**UNIVERSITAS
MIKROSKIL**

**INFORMATION TECHNOLOGY MASTER PROGRAM
SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN
KOMPUTER MIKROSKIL**

MEDAN

2020

LEMBAR PENGESAHAN
PENINGKATAN KINERJA KLASIFIKASI GENETIC
ALGORITHM MELALUI PENANGANAN INCOMPLETE
DATA DAN REDUKSI

(Studi Kasus : Dataset Faktor Risiko Kanker Serviks)

TESIS

Diajukan Untuk Melengkapi Persyaratan Guna
Mendapatkan Gelar Magister
Program Studi Teknologi Informasi

Oleh :

SRI NOVIDA SARI
NIM. 174211054

Disetujui Oleh:

BOSEN PEMBIMBING I,


Dr. Pahala Sirait, S. Kom., M.Kom

Medan, 30 Januari 2020

Diketahui dan Disahkan Oleh:

Ketua Program Studi
Magister Teknologi Informasi


Dr. Ronsen Purba, M.Sc

LEMBAR PERNYATAAN

Saya yang membuat pernyataan ini adalah mahasiswa/I Jurusan/Program Studi Magister Teknologi Informasi STMIK MIKROSKIL Medan dengan identitas mahasiswa sebagai berikut :

Nama : SRI NOVIDA SARI
NIM : 174211054

Saya telah melaksanakan penelitian dan penulisan Tugas Akhir/Tesis dengan judul "PENINGKATAN KINERJA KLASIFIKASI ALGORITMA GENETIKA MELALUI PENANGANAN INCOMPLETE DATA DAN REDUKSI" (Studi Kasus : Dataset Faktor Risiko Kanker Serviks) , dengan ini saya menyatakan sebenar-benarnya bahwa penelitian dan penulisan Tugas Akhir/Tesis tersebut merupakan hasil karya saya sendiri (tidak menyuruh orang lain yang mengerjakannya) dan semua sumber, baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar. Bila di kemudian hari ternyata terbukti bahwa bukan saya yang mengerjakannya (membuatnya), maka saya bersedia dikenakan sanksi yang telah ditetapkan oleh STMIK Mikroskil Medan, yakni Pencabutan Ijazah yang telah saya terima dan ijazah tersebut dinyatakan tidak sah.

Selain itu, demi pengembangan ilmu pengetahuan, saya menyetujui untuk memberikan kepada STMIK Mikroskil Medan Hak Bebas Royalti Non-eksklusif (Non-exclusive Royalty Free Right) atas Tugas Akhir/Tesis saya beserta perangkat yang ada (jika diperlukan). Dengan hak ini, STMIK Mikroskil Medan berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data(database), merawat, dan mempublikasikan Tugas Akhir/Tesis saya, secara keseluruhan atau hanya sebagian atau hanya ringkasannya saja dalam bentuk format tercetak maupun elektronik, selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik hak cipta. Menyatakan juga bahwa saya akan mempertahankan hak eksklusif saya untuk menggunakan seluruh atau sebagian isi Tugas Akhir/ Tesis saya guna pengembangan karya dimasa depan , misalnya dalam bentuk artikel, buku, ataupun perangkat lunak.

Demikian pernyataan ini saya perbuat dengan sungguh-sungguh, dalam keadaan sadar dan tanpa ada tekanan dari pihak manapun.

Medan, 25 Januari 2020

Saya yang membuat pernyataan,



Sri Novida Sari
SRI NOVIDA SARI

ABSTRAK

Kanker serviks adalah keganasan ginekologis terkemuka di seluruh dunia. penelitian ini menyajikan teknik optimasi dan menunjukkan pemilihan fitur untuk kombinasi atribut terbaik dalam resiko penyakit kanker serviks. Ada tiga puluh dua atribut dengan delapan ratus lima puluh delapan sampel. Selain itu, data ini juga memiliki nilai yang tidak lengkap karena masalah privasi responden dan data ketidakseimbangan. Oleh karena itu, digunakan Teknik preprocessing dengan metode regression imputation dan normalisasi data. Selanjutnya, teknik reduksi atribut dengan metode optimum index factor (OIF) diperlukan untuk meningkatkan akurasi hasil. Hasil penelitian menunjukkan bahwa perbandingan optimasi dengan reduksi dan tanpa reduksi atribut memiliki selisih 1%. optimasi dengan tanpa reduksi atribut di dapat hasil 95% sedangkan hasil dengan reduksi atribut adalah 94%.

Kata Kunci : *Preprocessing, Incomplete data*, normalisasi data, *optimum index factor*, Optimasi



UNIVERSITAS
MIKROSKIL

ABSTRACT

Cervical cancer is a leading gynecologic malignancy worldwide. this research presents classification techniques and shows the selection of features for the best combination of attributes in the risk of cervical cancer. There are thirty two attributes with eight hundred fifty eight samples. In addition, this data also has incomplete values due to respondents' privacy issues and imbalance data. Therefore, preprocessing techniques are used with regression imputation and data normalization methods. Furthermore, attribute reduction techniques with the optimum index factor (OIF) method are needed to improve the accuracy of the results. The results showed that the comparison of optimization with reduction and without attribute reduction has a difference of 1%. optimization with no attribute reduction results in 95% while results with attribute reduction are 94%.

Keywords: Preprocessing, Incomplete data, data normalization, optimum index factor, optimization.



UNIVERSITAS
MIKROSKIL

KATA PENGANTAR

Puji dan syukur penulis panjatkan kepada Tuhan Yang Maha Esa karena atas berkat dan rahmat-Nya, penulis dapat menyelesaikan Tesis yang berjudul “*Optimasi genetic algorithm melalui penanganan incomplete data dan reduksi*”.

Tesis ini dibuat untuk melengkapi persyaratan kurikulum pada Program Studi Magister Teknologi Informasi Strata Dua, STMIK Mikroskil Medan.

Penulis menyampaikan banyak terima kasih kepada:

1. Bapak Dr. Pahala Sirait, S.T., M.Kom., selaku Dosen Pembimbing dan Ketua STMIK Mikroskil Medan yang telah membimbing penulis dalam penyelesaian tesis ini.
2. Bapak Arwin Halim S.Kom., M.Kom., selaku Dosen Pendamping Pembimbing yang telah membimbing penulis dalam proses penyelesaian tesis ini.
3. Bapak Dr. Ronsen Purba, M.Sc., selaku Ketua Program Studi Magister Teknologi Informasi STMIK Mikroskil Medan.
4. Bapak atau Ibu Dosen STMIK Mikroskil Medan yang telah membantu proses penulisan tesis ini.
5. Anggota keluarga, suami, teman dan semua pihak yang terus memberikan dukungan penuh kepada penulis selama proses penulisan tesis ini.

Penulis menyadari bahwa masih terdapat banyak kekurangan yang ada. Oleh sebab itu, kritik dan saran yang bersifat membangun akan sangat diterima. Akhir kata, semoga Tesis ini dapat bermanfaat bagi masyarakat. Terima kasih.

Medan, 30 Januari 2020

Penulis

DAFTAR ISI

ABSTRAK	i
ABSTRACT	ii
KATA PENGANTAR	iii
DAFTAR ISI.....	iv
DAFTAR GAMBAR	vi
DAFTAR TABEL.....	vii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Masalah Penelitian	3
1.2.1 Identifikasi Masalah	3
1.2.2 Rumusan Masalah.....	3
1.3 Tujuan Dan Manfaat Penelitian	3
1.4 Batasan Masalah.....	4
1.5 Metodologi Penelitian	4
1.6 Sistematika Penelitian	5
BAB II KAJIAN LITERATUR.....	7
2.1 Tinjauan Pustaka.....	7
2.1.1 Kanker Serviks	8
2.1.2 Machine Learning	14
2.1.3 Incomplete data	16
2.1.4 Reduksi.....	18
2.1.4 Genetic Algorithm	19
2.2 Kerangka Konsep / Pola Pikir Pemecahan Masalah	31
BAB III METODOLOGI PENELITIAN.....	33
3.1 Analisis Masalah	33
3.2 Metode Penelitian.....	33
3.2.1 Incomplete data	35
3.2.1 Normalisasi Min-Max Method.....	37

3.2.2	Optimum Index Factor (OIF).....	38
3.2.3	Optimasi Genetic Algorithm.....	41
3.3	Data yang Digunakan.....	45
3.4	Alat – Alat Penelitian.....	45
3.5	Metode Analisis.....	46
3.6	Jadwal Penelitian.....	48
BAB IV HASIL DAN PEMBAHASAN.....		49
4.1	Hasil.....	49
4.1.1	Pembagian Data.....	49
4.1.2	Imputasi incomplete data dengan <i>Regresi Imputation</i>	51
4.1.3	Normalisasi Min-Max Method.....	52
4.1.4	Hasil Pengujian Reduksi Atribut.....	54
4.1.5	Hasil Pengujian optimasi dengan Genetic Algorithm.....	61
4.2	Pembahasan.....	62
4.2.1	Incomplete data dan Normalisasi.....	62
4.2.2	Reduksi Atribut dengan <i>Optimum Index Factor (OIF)</i>	63
4.2.3	Optimasi <i>Genetic Algorithm</i>	63
BAB V PENUTUP.....		65
5.1	Kesimpulan.....	65
5.2	Saran.....	65
DAFTAR PUSTAKA.....		66

UNIVERSITAS
MIKROSKIL

DAFTAR GAMBAR

Gambar 2-1 Diagram Alir <i>Genetic Algorithm</i>	34
Gambar 3-1 Metode Penelitian	34
Gambar 3-2 <i>Incomplete Regression Imputation</i>	35
Gambar 3-3 Normalisasi Data.....	38
Gambar 3-4 Flowchart <i>Optimum index Factor (OIF)</i>	39
Gambar 3-5 <i>Genetic Algorithm</i>	42
Gambar 4-1 Plot <i>Incomplete data</i>	51
Gambar 4-2 Sebelum imputasi pada <i>incomplete data</i>	52
Gambar 4-3 Sesudah Imputasi dengan <i>Regression Imputation</i>	52



UNIVERSITAS
MIKROSKIL

DAFTAR TABEL

Tabel 2-1 Contoh Crossover 1-titik	26
Tabel 2-2 Contoh Crossover 2-titik	26
Tabel 2-3 Contoh Crossover Seragam	27
Tabel 2-4 Contoh Mutasi Pada Pengkodean biner	28
Tabel 2-5 Contoh Mutasi Pada Pengkodean Permutasi	28
Tabel 3-1 Bentuk Confusion Matrix Dari Dua Kelas	47
Tabel 3-2 Jadwal Penelitian	48
Tabel 4-1 Informasi Dataset.....	50
Tabel 4-2 Data Sebelum Normalisasi	53
Tabel 4-3 Data Sesudah Normalisasi.....	53
Tabel 4-4 Reduksi 1 Atribut.....	54
Tabel 4-5 Reduksi 2 Atribut.....	54
Tabel 4-6 Reduksi 3 Atribut.....	55
Tabel 4-7 Reduksi 4 Atribut.....	56
Tabel 4-8 Reduksi 5 Atribut.....	56
Tabel 4-9 Reduksi 6 Atribut.....	57
Tabel 4-10 Reduksi 7 Atribut.....	57
Tabel 4-11 Reduksi 8 Atribut.....	58
Tabel 4-12 Reduksi 9 Atribut.....	58
Tabel 4-13 Reduksi 10 Atribut.....	59
Tabel 4-14 Reduksi 11 Atribut.....	60
Tabel 4-15 Reduksi 12 Atribut.....	60
Tabel 4-16 Pengujian Optimasi GA tanpa reduksi atribut	61
Tabel 4-17 Pengujian Optimasi GA dengan reduksi atribut	62
Tabel 4-18 Kombinasi OIF	63
Tabel 4-19 Hasil Optimasi Genetic Algorithm	64