

**REDUKSI ATRIBUT PADA DATASET KREDIT DENGAN
ALGORITMA GENETIKA DAN KLASIFIKASI DENGAN
METODE K-NN**

TESIS

Oleh:

Kent

NIM. 174212101



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

MIKROSKIL

MEDAN

2020

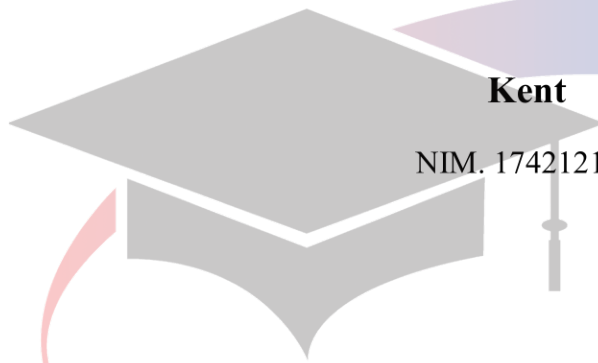
**ATTRIBUTE REDUCTION IN CREDIT DATASETS WITH
GENETIC ALGORITHM AND CLASSIFICATION USE K-NN**

THESIS

By:

Kent

NIM. 174212101



**UNIVERSITAS
MIKROSKIL**

**INFORMATION TECHNOLOGY MASTER'S PROGRAM
SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER**

MIKROSKIL

MEDAN

2020

LEMBAR PERNYATAAN

Saya yang membuat pernyataan ini adalah mahasiswa Jurusan/ Program Studi S-2 Magister Teknologi Informasi STMIK Mikroskil Medan dengan identitas mahasiswa sebagai berikut:

Nama : Kent
NIM : 174212101
Peminatan : Teknologi Informasi

Saya telah melaksanakan penelitian dan penulisan tesis dengan judul "REDUKSI ATRIBUT PADA DATASET KREDIT DENGAN ALGORITMA GENETIKA DAN KLASIFIKASI DENGAN METODE K-NN". Dengan ini saya menyatakan dengan sebenar- benarnya bahwa penelitian dan penulisan 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 tesis saya. Dengan hak ini, STMIK Mikroskil Medan berhak menyimpan, mengalih media/ formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan tesis saya, secara keseluruhan atau hanya sebagian atau hanya ringkasannya saja dalam bentuk format tercetak dan/ atau 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 tesis saya guna pengembangan karya di masa 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

Saya yang membuat pernyataan



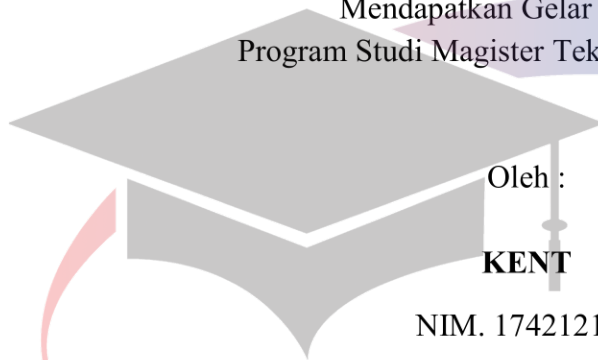
Kent

LEMBARAN PENGESAHAN

REDUKSI ATRIBUT PADA DATASET KREDIT DENGAN ALGORITMA GENETIKA DAN KLASIFIKASI DENGAN METODE K-NN

TESIS

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



Oleh :

KENT

NIM. 174212101

Disetujui Oleh:

Dosen Pembimbing I,

UNIVERSITAS
MIKROSKIL

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

Medan, 24 Juli 2020

Diketahui dan Disahkan Oleh:

Ketua Program Studi
Magister Teknologi Informasi,

Dr. Ronsen Purba, M.Sc.

ABSTRAK

Reduksi atribut ialah proses pengurangan ataupun penghilangan atribut untuk meningkatkan tingkat efisiensi pengambilan keputusan dan juga menjaga klasifikasi terhadap objek penelitian. Masalah utama dalam penelitian ialah analisis data yang kompleks dan penambangan dalam jumlah besar butuh waktu yang lama dan proses yang panjang, sehingga membuat analisis tersebut tidak praktis atau tidak layak. Jumlah atribut yang terlalu banyak juga dapat mengakibatkan proses penarikan kesimpulan yang terlalu panjang. Proses reduksi atribut digunakan untuk menentukan kombinasi atribut yang tepat dengan cara menajamkan, membuang yang kurang diperlukan, sehingga atribut yang dilibatkan adalah atribut- atribut dengan pengaruh paling tinggi saja, namun hasilnya tetap optimal. Pada penelitian ini metode yang digunakan untuk mereduksi atribut ialah algoritma genetika dan untuk tahap klasifikasinya menggunakan metode *K-NN*. Pengujian dilakukan terhadap *dataset* sekunder. Hasil yang diperoleh ialah setelah proses reduksi atribut dilakukan ialah tetap akan memberikan hasil akurasi yang mirip, ataupun jika ada pengurangan tidak akan terlalu berarti. Perbandingan sebelum dan sesudah pengurangan atribut membuktikan bahwa optimasi hasil tetap terjaga.

Kata Kunci: Reduksi Atribut, Algoritma Genetika, Klasifikasi, *K-NN*

ABSTRACT

Attribute reduction is the process of reducing or removing attributes to increase the efficiency level of decision making and also maintaining the classification of the research object. The main problem with research is that the analysis of complex data and mining of large amounts takes a long time and is a long process, making such analysis impractical or not feasible. Too many attributes can also lead to a lengthy process of drawing conclusions. The process of attribute reduction is used to determine the right combination of attributes by sharpening, removing those that are less needed, so that the attributes involved are the attributes with the highest effect, but the results are still optimal. In this research, the method used to reduce attributes is genetic algorithm and for the classification stage using the K-NN method. Tests are carried out on the secondary dataset. The results obtained are that after the attribute reduction process is carried out, it will still provide similar accuracy results, or if there is a reduction it will not be too meaningful. Comparison before and after attribute reduction proves that the result optimization is maintained.

Keywords: Attribute Reduction, Genetic Algorithm, Classification, K-NN

UNIVERSITAS
MIKROSKIL

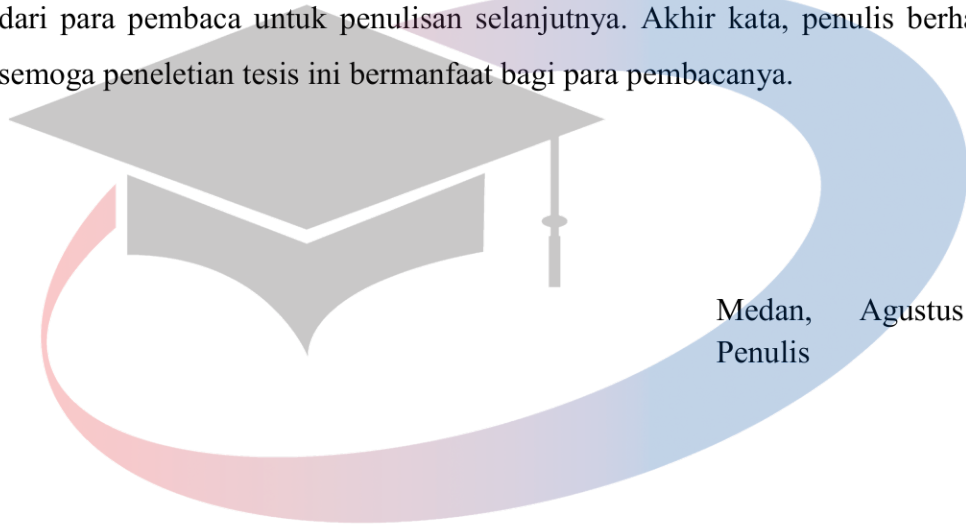
KATA PENGANTAR

Puji serta syukur penulis ucapkan kepada Tuhan Yang Maha Esa atas berkat rahmat dan karuniaNya yang telah memberikan kemampuan, kebijaksanaan, kesehatan dan kesempatan sehingga dapat menyelesaikan penelitian tesis yang berjudul **“REDUKSI ATRIBUT PADA DATASET KREDIT DENGAN ALGORITMA GENETIKA DAN KLASIFIKASI DENGAN METODE K-NN”**. Tesis ini merupakan salah satu syarat untuk mendapatkan gelar Magister Teknologi Informasi (M.T.I) pada program studi magister Sekolah Tinggi Manajemen Informatika dan Komputer Mikroskil Medan. Dalam penulisan tesis ini penulis banyak menerima bantuan, bimbingan, nasehat, dukungan, dan dorongan semangat dari berbagai pihak. Oleh karena itu dengan hati yang tulus penulis ingin mengucapkan terimakasih kepada :

1. Bapak Dr. Pahala Sirait, S.T., M.Kom. selaku Dosen Pembimbing dan Ketua Sekolah Tinggi Manajemen Informatika dan Komputer Mikroskil Medan yang telah memberikan waktunya untuk memberikan bimbingan dalam menyelesaikan laporan tesis ini.
2. Bapak Arwin Halim, S.Kom., M.Kom., selaku Dosen Pendamping Pembimbing yang telah memberikan waktunya untuk memberikan bimbingan dalam menyelesaikan laporan tesis ini.
3. Bapak Dr. Ronsen Purba, M.Sc. selaku Ketua Program Studi Magister Teknologi Informasi.
4. Bapak Djoni, S.Kom., M.T.I., selaku Wakil Ketua I Sekolah Tinggi Manajemen Informatika dan Komputer Mikroskil Medan.
5. Teristimewa kepada kedua orang tua tercinta yang tidak pernah berkata lelah dan selalu membantu dan mendukung penulis. Terima kasih penulis ucapkan untuk kasih sayang, doa, pengorbanan, serta dukungan yang selama ini diberikan.

6. Buat sahabat– sahabatku dan kepada teman-teman yang banyak mendukung, membantu dan memberi saran dalam pembuatan laporan tesis ini.
7. Seluruh Dosen dan staff Pengajar yang telah membantu mendidik dan memberikan ilmu kepada penulis selama berada di bangku kuliah.

Penulis menyadari penelitian tesis ini masih jauh dari sempurna. Oleh karena itu, saran dan kritik yang bersifat membangun sangat penulis harapkan dari para pembaca untuk penulisan selanjutnya. Akhir kata, penulis berharap semoga penelitian tesis ini bermanfaat bagi para pembacanya.



Medan, Agustus 2020
Penulis

Kent
174212101
UNIVERSITAS
MIKROSKIL

DAFTAR ISI

ABSTRAK	i
ABSTRACT	ii
KATA PENGANTAR	iii
DAFTAR ISI	v
DAFTAR GAMBAR	viii
DAFTAR TABEL	ix
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Masalah Penelitian	2
1.2.1 Identifikasi Masalah	2
1.2.2 Rumusan Masalah	3
1.3 Tujuan dan Manfaat Penelitian	3
1.4 Batasan Masalah	3
1.5 Metodologi Penelitian	4
1.6 Sistematika Penulisan	4
BAB II KAJIAN LITERATUR	6
2.1 Tinjauan Pustaka	6
2.1.1 Kredit	6
2.1.2 Data Mining	7
2.1.2.1 Kegunaan Data Mining	8
2.1.2.2 Aplikasi- Aplikasi <i>Data Mining</i>	9
2.1.2.3 Proses Data Mining	12
2.1.3 Reduksi Atribut	12

2.1.3.1	Strategi Reduksi Atribut.....	13
2.1.3.2	Seleksi Atribut.....	13
2.1.4	Algoritma Genetika.....	16
2.1.4.1	Pemetaan dari Proses Alamiah ke Proses Komputasi	18
2.1.4.2	Struktur Data Algoritma Genetika	20
2.1.4.3	Seleksi dalam Algoritma Genetika.....	21
2.1.4.4	Penyilangan dalam Algoritma Genetika	24
2.1.4.5	Mutasi dalam Algoritma Genetika	27
2.1.5	KNN.....	28
2.1.6	Penelitian Terkait Sebelumnya	29
2.2	Kerangka Konsep/ Pola Pikir Pemecahan Masalah	30
BAB III METODOLOGI PENELITIAN		32
3.1	Analisis Masalah	32
3.2	Metode Penelitian.....	33
3.2.1	Prapemrosesan Data.....	34
3.2.2	Reduksi Atribut.....	36
3.2.3	Proses Klasifikasi <i>K-Nearest Neighbor</i>	46
3.3	Metode Pengumpulan Data	52
3.4	Alat- Alat Penelitian.....	52
3.5	Teknik Analisis Data.....	52
BAB IV HASIL DAN PEMBAHASAN		54
4.1	Hasil	54
4.1.1	Hasil Reduksi Atribut	54
4.1.2	Hasil Klasifikasi.....	58
4.2	Pembahasan.....	68

BAB V PENUTUP	70
5.1 Kesimpulan	70
5.2 Saran.....	70
DAFTAR PUSTAKA	71



UNIVERSITAS MIKROSKIL

DAFTAR GAMBAR

Gambar II-1. Kerangka Kerja Penerapan Algoritma Genetika untuk Menyelesaikan Suatu Masalah Optimasi (Zukhri, Z., 2014)	21
Gambar II-2. Ilustrasi Seleksi dengan Model Roda Rolet (Zukhri, Z., 2014)	22
Gambar II-3. Ilustrasi Metode Penyilangan n-Titik (Zukhri, Z., 2014)	26
Gambar II-4. Ilustrasi Metode Penyilangan Seragam (Zukhri, Z., 2014)	26
Gambar II-5. Ilustrasi Metode Penyilangan Berbasis Posisi (Zukhri, Z., 2014)...	27
Gambar II-6. Ilustrasi Metode Penyilangan Berbasis Urutan (Zukhri, Z., 2014) .	27
Gambar II-7. Ilustrasi Metode Mutasi untuk Kode Biner (Zukhri, Z., 2014)	28
Gambar II-8. Klasifikasi Dua Kelas Menggunakan kNN dengan $k = 3$ (Suyanto, 2019)	28
Gambar II-9. Kerangka Konsep Pemecahan Masalah	31
Gambar III-1. Metode Penelitian	33
Gambar III-2. Langkah Algoritma Genetika.....	36
Gambar III-3. Langkah Algoritma K-Nearest Neighbor.....	51
Gambar IV-1. Hasil Akurasi Tanpa Reduksi Atribut.....	60
Gambar IV-2. Hasil Akurasi Setelah Reduksi 1 Atribut	64
Gambar IV-3. Hasil Akurasi Setelah Reduksi 2 Atribut.....	64
Gambar IV-4. Hasil Akurasi Setelah Reduksi 5 Atribut	65
Gambar IV-5. Hasil Akurasi Setelah Reduksi 10 Atribut.....	65
Gambar IV-6. Perbandingan Hasil Akurasi Antara <i>Dataset</i> Tanpa Reduksi Dengan <i>Dataset</i> Setelah Direduksi	66
Gambar IV-7. Perbandingan Hasil Akurasi Antara <i>Dataset</i> Tanpa Reduksi Dengan <i>Dataset</i> Setelah Direduksi 10 Atribut.....	67

DAFTAR TABEL

Tabel II-1. Data Pelanggan Telepon Seluler	14
Tabel II-2. Contoh Data Atribut.....	15
Tabel II-3. Contoh Data Atribut yang Sudah Direduksi	16
Tabel III-1. Karakter Masing- Masing Atribut Dataset German Credit Data	34
Tabel III-2. Pembentukan Kombinasi Atribut	37
Tabel III-3. Memisahkan Record Atribut Pertama Sesuai dengan Class.....	38
Tabel III-4. Memisahkan Record Atribut Kedua Sesuai dengan Class	38
Tabel III-5. Memisahkan Record Atribut Ketiga Sesuai dengan Class	38
Tabel III-6. Memisahkan Record Atribut Keempat Sesuai dengan Class	39
Tabel III-7. Memisahkan Record Atribut Kelima Sesuai dengan Class	39
Tabel III-8. Memisahkan Record Atribut Keenam Sesuai dengan Class.....	40
Tabel III-9. Memisahkan Record Atribut Ketujuh Sesuai dengan Class.....	40
Tabel III-10. Memisahkan Record Atribut Kedelapan Sesuai dengan Class.....	40
Tabel III-11. Memisahkan Record Atribut Kesembilan Sesuai dengan Class.....	41
Tabel III-12. Memisahkan Record Atribut Kesepuluh Sesuai dengan Class.....	41
Tabel III-13. Memisahkan Record Atribut Kesebelas Sesuai dengan Class.....	41
Tabel III-14. Memisahkan Record Atribut Kedua Belas Sesuai dengan Class.....	42
Tabel III-15. Memisahkan Record Atribut Ketiga Belas Sesuai dengan Class ...	42
Tabel III-16. Memisahkan Record Atribut Keempat Belas Sesuai dengan Class.	42
Tabel III-17. Memisahkan Record Atribut Kelima Belas Sesuai dengan Class ...	42
Tabel III-18. Memisahkan Record Atribut Keenam Belas Sesuai dengan Class..	43
Tabel III-19. Memisahkan Record Atribut Ketujuh Belas Sesuai dengan Class ..	43
Tabel III-20. Memisahkan Record Atribut Kedelapan Belas Sesuai dengan Class	43
Tabel III-21. Memisahkan Record Atribut Kesembilan Belas Sesuai dengan Class	43
Tabel III-22. Memisahkan Record Atribut Kedua Puluh Sesuai dengan Class	44
Tabel III-23. Pembentukan Individu	44
Tabel III-24. Contoh Data Training	47

Tabel III-25. Data Testing.....	47
Tabel III-26. Hasil Euclidean Distance.....	50
Tabel III-27. Hasil Pengurutan Dataset Awal.....	50
Tabel IV-1. Daftar Atribut Sebelum Dilakukan Reduksi Atribut.....	55
Tabel IV-2. Daftar Atribut Setelah Dilakukan Reduksi Atribut.....	56
Tabel IV-3. Perbandingan Pencarian Individu dengan Fitness Tertinggi.....	58
Tabel IV-4. Hasil Klasifikasi Tanpa Reduksi Atribut.....	59
Tabel IV-5. Hasil Klasifikasi Setelah Reduksi 1 Atribut.....	61
Tabel IV-6. Hasil Klasifikasi Setelah Reduksi 2 Atribut.....	61
Tabel IV-7. Hasil Klasifikasi Setelah Reduksi 5 Atribut.....	62
Tabel IV-8. Hasil Klasifikasi Setelah Reduksi 10 Atribut.....	62
Tabel IV-9. Hasil Akurasi Tertinggi Tanpa Proses Reduksi Atribut.....	67
Tabel IV-10. Hasil Akurasi Tertinggi Setelah Proses Reduksi Atribut.....	67



UNIVERSITAS MIKROSKIL