

**Model Data Mining Dalam Menganalisis Faktor-Faktor Yang
Mempengaruhi Kemiskinan di Sumatera Selatan**



TESIS

Hardianto

Enterprise IT Infrastructure

212420024

PROGRAM STUDI TEKNIK INFORMATIKA S-2

PROGRAM PASCASARJANA

UNIVERSITAS BINA DARMA

PALEMBANG

2024

**Model Data Mining Dalam Menganalisis Faktor-Faktor Yang
Mempengaruhi Kemiskinan di Sumatera Selatan**

**Tesis ini diajukan sebagai salah satu syarat
untuk memperoleh gelar**

MAGISTER KOMPUTER



Hardianto

Enterprise IT Infrastructure

212420024

PROGRAM STUDI TEKNIK INFORMATIKA S-2

PROGRAM PASCASARJANA

UNIVERSITAS BINA DARMA

PALEMBANG

2024

Halaman Pengesahan Pembimbing Tesis

Judul Tesis : MODEL DATA MINING DALAM MENGANALISIS
FAKTOR-FAKTOR YANG MEMPENGARUHI
KEMISKINAN DI SUMATERA SELATAN

Oleh HARDIANTO NIM 212420024 Tesis ini telah disetujui dan disahkan oleh Tim
Penguji Program Studi Teknik Informatika – S2 Konsentrasi ENTERPRISE IT
INFRASTRUKTUR, Program Pascasarjana Universitas Bina Darma pada 29
Agustus 2024 dan telah dinyatakan LULUS.

Palembang, 29 Agustus 2024

Mengetahui,

Program Studi Teknik Informatika – S2

Universitas Bina Darma

Ketua,



.....
Dr. Usman Ependi, M.Kom

Pembimbing,

.....
Dr. Yesi Novaria Kunang, S.T., M.Kom

Halaman Pengesahan Penguji Tesis

Judul Tesis : MODEL DATA MINING DALAM MENGANALISIS
FAKTOR-FAKTOR YANG MEMPENGARUHI
KEMISKINAN DI SUMATERA SELATAN

Oleh HARDIANTO NIM 212420024 Tesis ini telah disetujui dan disahkan oleh Tim Penguji Program Studi Teknik Informatika – S2 Konsentrasi ENTERPRISE IT INFRASTRUKTUR, Program Pascasarjana Universitas Bina Darma pada 29 Agustus 2024 dan telah dinyatakan LULUS.

Palembung, 29 Agustus 2024

Mengetahui,

Program Pascasarjana
Universitas Bina Darma
Direktur,

Tim Penguji:

Penguji I,



Prof. Dr. Ir. Achmad Syarifudin, M.Sc.

Dr. Yesi Novaria Kunang, S.T., M.Kom

Penguji II,

Prof. Dr. Edi Surya Negara, M.Kom

Penguji III,

Dr. Tata Sutabri, S.Kom., M.M.S.I

SURAT PERNYATAAN

Saya yang bertanda tangan dibawah ini:

Nama : HARDIANTO
NIM : 212420024

Dengan ini menyatakan bahwa:

1. Karya tulis Saya Tesis ini adalah asli dan belum pernah diajukan untuk mendapatkan gelar akademik Magister di Universitas Bina Darma;
2. Karya tulis ini murni gagasan, rumusan dan penelitian Saya sendiri dengan arahan tim pembimbing;
3. Dalam karya tulis ini tidak terdapat karya atau pendapat yang telah ditulis atau dipublikasikan orang lain, kecuali secara tertulis dengan jelas dikutip dengan mencantumkan nama pengarang dan memasukkan ke dalam daftar pustaka;
4. Karena yakin dengan keaslian karya tulis ini, Saya menyatakan bersedia Tesis yang Saya hasilkan di unggah ke internet;
5. Surat Pernyataan ini Saya tulis dengan sungguh-sungguh dan apabila terdapat penyimpangan atau ketidakbenaran dalam pernyataan ini, maka Saya bersedia menerima sanksi dengan aturan yang berlaku di perguruan tinggi ini.

Demikian Surat Pernyataan ini saya buat agar dapat dipergunakan sebagaimana mestinya.

Palembang, 29 Agustus 2024
Yang Membuat Pernyataan,



HARDIANTO
NIM : 212420024

Abstrak

Pengentasan kemiskinan merupakan prioritas utama dalam agenda pembangunan nasional Indonesia. Berbagai program perlindungan sosial telah diluncurkan untuk mengurangi angka kemiskinan. Presiden Joko Widodo menargetkan pengentasan kemiskinan ekstrem pada tahun 2024 melalui Inpres Nomor 4 Tahun 2022. Namun demikian, tantangan masih tetap ada, khususnya di daerah seperti Sumatera Selatan, dimana tingkat kemiskinan masih tinggi meskipun pendapatannya relatif tinggi. Karena faktor sosio-ekonomi yang kompleks merupakan faktor penentu utama kemiskinan, diperlukan pendekatan analitis yang lebih holistik.

Penelitian ini bertujuan untuk menganalisis faktor-faktor yang mempengaruhi kemiskinan di Sumatera Selatan melalui pendekatan multidimensi dengan menggunakan teknik *data mining* dan visualisasi data. Fokus penelitiannya adalah mengidentifikasi faktor-faktor yang berkontribusi signifikan terhadap kemiskinan dan memvisualisasikan pola dan tren dalam data. Data yang digunakan berasal dari Survei Sosial Ekonomi Nasional (Susenas) Maret 2023 yang dilakukan Badan Pusat Statistik (BPS). Metode penelitian yang digunakan meliputi perencanaan proyek, persiapan data, analisis data, dan visualisasi hasil. Penelitian ini terkait dengan penelitian sebelumnya yang dilakukan di bidang kemiskinan dan teknik analisis data di Indonesia. Hasil penelitian diharapkan dapat memberikan wawasan yang lebih mendalam mengenai faktor-faktor yang mempengaruhi kemiskinan di Sumatera Selatan dan berkontribusi pada pengembangan kebijakan yang lebih efektif untuk pengentasan kemiskinan di wilayah tersebut.

Kata Kunci: *Data Mining*, kemiskinan, analisis multidimensi, visualisasi data, susenas, Sumatera selatan

Abstract

Poverty alleviation is a top priority in Indonesia's national development agenda. Various social protection programs have been launched to reduce poverty rates. President Joko Widodo has set a target of eradicating extreme poverty by 2024 through Presidential Instruction No. 4 of 2022. However, challenges still remain, particularly in regions such as South Sumatra, where poverty rates remain high despite relatively high incomes. Given the complex socio-economic factors as the main determinants of poverty, a more holistic analytical approach is necessary.

This study aims to analyze the factors influencing poverty in South Sumatra through a multidimensional approach using data mining techniques and data visualization. The research focus is to identify the factors that contribute significantly to poverty and to visualize patterns and trends in the data. The data used is from the 2023 March National Socio-Economic Survey (Susenas) conducted by the Central Statistics Agency (BPS). The research methods include project planning, data preparation, data analysis, and visualization of results. This research is related to previous studies conducted in the field of poverty and data analysis techniques in Indonesia. The research results are expected to provide deeper insights into the factors influencing poverty in South Sumatra and contribute to the development of more effective policies for poverty alleviation in the region.

Keywords: *Data Mining, Poverty, Multidimensional Analysis, Data Visualization*

MOTTO DAN HALAMAN PERSEMBAHAN

اعْمَلْ لِدُنْيَاكَ كَأَنَّكَ تَعِيشُ أَبَدًا، وَاعْمَلْ لِآخِرَتِكَ كَأَنَّكَ تَمُوتُ غَدًا

“Bekerjalah untuk duniamu seakan-akan engkau akan hidup selamanya. Dan bekerjalah untuk akhiratmu seakan-akan engkau akan mati besok pagi.”

Alhamdulillah, segala puji syukur kepada Allah SWT tesis ini pada akhirnya dapat kami selesaikan. Tesis ini saya persembahkan untuk:

1. Orang tua yang telah mendoakan setiap langkah saya,
2. Istriku dan anak-anakku atas segala usaha dan pengorbanan yang telah dilakukan hingga terselesainya tesis ini.

KATA PENGANTAR

Segala puji bagi Allah SWT atas limpahan rahmat dan ridho-Nya sehingga penulis dapat menyelesaikan tesis yang berjudul “Model Data Mining dalam Menganalisis Faktor-Faktor yang mempengaruhi Kemiskinan di Sumatera Selatan”. Tesis ini disusun sebagai salah satu syarat untuk memperoleh gelar Magister Komputer pada Universitas Bina Darma Palembang. Penulis mengucapkan terima kasih atas bantuannya kepada:

1. Prof. Dr. Sunda Ariana, M.Pd., M.M. selaku Rektor Universitas Bina Darma.
2. Prof. Dr. Ir. Achmad Syarifudin, M.Sc., selaku Direktur Pascasarjana Universitas Bina Darma
3. Dr. Usman Ependi, M.Kom, selaku Ketua Program Studi Magister Teknik Informatika Universitas Bina Darma.
4. Dr. Yesi Novaria Kunang, S.T., M.Kom, selaku dosen pembimbing yang telah meluangkan banyak waktu, tenaga, serta pikiran dalam membimbing.
5. Kepala BPS Provinsi Sumatera Selatan selaku pimpinan tempat bekerja atas dukungannya dalam menyelesaikan kuliah kami.
6. Teman-teman kerja khususnya Muhammad Kharis dan Nugraha Pukuh atas bantuannya dalam menyelesaikan tesis.
7. Keluarga besar penulis khususnya orang tua, istri dan anak-anak penulis atas perhatian dan pengorbanan yang telah dilakukan hingga tesis ini dapat terselesaikan.
8. Serta pihak-pihak lain yang tidak mungkin disebutkan disini atas bantuannya.

Penulis sadar penyusunan tesis ini masih jauh dari kesempurnaan, oleh karena itu kritik serta saran sangat penulis harapkan untuk perbaikan selanjutnya. Akhirnya, penulis berharap semoga tesis ini bermanfaat bagi semua pihak.

Palembang, September 2024

Penulis

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PENGESAHAN	iii
SURAT PERNYATAAN	v
ABSTRAK	vi
MOTTO DAN HALAMAN PERSEMBAHAN	viii
KATA PENGANTAR	xi
DAFTAR ISI	x
DAFTAR GAMBAR	xi
DAFTAR TABEL	xii
BAB I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	4
1.3. Batasan Masalah	4
1.4. Tujuan Penelitian	5
1.5. Manfaat Penelitian	5
1.6. Sistematika Penulisan	6
BAB II LANDASAN/KERANGKA PEMIKIRAN	7
2.1. Tinjauan Pustaka	7
2.1.1 Pentingnya Pengukuran Kemiskinan	7
2.1.2 Konsep Dasar dan Definisi Kemiskinan	8
2.1.3 <i>Data Mining</i>	14
2.1.4 Data Visualization	16
2.1.5 Python	18
2.2. Penelitian Terdahulu	20
2.3. Kerangka Pemikiran	22
BAB III METODOLOGI PENELITIAN	23
3.1. Dasar Penelitian	23
3.2. Ruang Lingkup Penelitian	24
3.3. Instrumen Penelitian	24
3.3.1 Data	24
3.3.2 Perangkat Lunak (<i>software</i>)	25
3.3.3 Perangkat Keras (<i>hardware</i>)	26
3.4. Metode Pengumpulan Data	26
3.5. Metode Pengolahan Data	27
3.5.1 <i>Data Selection</i>	27
3.5.2 <i>Preprocessing</i>	28
3.5.3 Pemilihan Algoritma (<i>Feature Selection</i>)	29
3.5.4 Permodelan Data (<i>Data Modelling</i>)	29
3.5.5 Evaluasi dan Interpretasi (<i>Evaluation and Interpretation</i>)	29
3.6. Jadwal Penelitian	29
BAB IV HASIL DAN PEMBAHASAN	30
4.1 <i>Data Selection</i>	30
4.2 <i>Preprocessing</i>	33
4.3 <i>Transformation</i>	34
4.4 <i>Data Mining</i>	38

4.5 <i>Interpretation/Evaluation</i>	41
BAB V PENUTUP	57
DAFTAR PUSTAKA	59
LAMPIRAN	62



DAFTAR GAMBAR

Gambar 1. Data mining sebagai proses KDD	14
Gambar 2. Kerangka berpikir	22
Gambar 3. Kerangka berpikir	23
Gambar 4. Pengecekan <i>missing value</i>	34
Gambar 5. <i>Heatmap</i> Perbandingan <i>performance</i> dari Model	40
Gambar 6. <i>feature importance</i> variabel dalam model	42
Gambar 7. PDP indeks aksesibilitas dan pengeluaran konsumsi makanan	46
Gambar 8. PDP luas lantai perkapita dan luas lantai tempat tinggal	46
Gambar 9. PDP pengeluaran konsumsi non makanan dan jumlah ART	47
Gambar 10. PDP umur dan indeks infrastruktur	47
Gambar 11. PDP status perkawinan dan indeks Pelayanan dasar	48
Gambar 12. PDP bahan bakar utama dan indeks kesulitan geografis 2021	48
Gambar 13. PDP status dalam usaha dan jumlah ART yang bekerja	49
Gambar 14. <i>Heatmap</i> Matriks Korelasi	53

DAFTAR TABEL

Tabel 1. Jadwal Penelitian	29
Tabel 2. Data Mikro Susenas Maret 2023	30
Tabel 3. Data Rumah Tangga Sampel Susenas Maret 2023	31
Tabel 4. Pengelompokan Variabel Data Susenas Maret 2023	35
Tabel 5. Kontribusi Variabel dalam analisis <i>feature importance</i>	42



