

**KLASIFIKASI PENYAKIT HIPERTENSI
PASIEN PUSKESMAS X KOTO I KABUPATEN TANAH DATAR
MENGUNAKAN METODE *DECISION TREE* ALGORITMA C4.5**

TUGAS AKHIR

*Diajukan sebagai salah satu persyaratan untuk memperoleh gelar
Ahli Madya*



Oleh

**ABID MAKARIM DYFA
NIM 19037002**

**PROGRAM STUDI DIPLOMA III STATISTIKA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS NEGERI PADANG
2022**

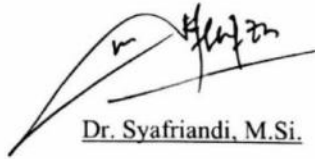
PERSETUJUAN UJIAN TUGAS AKHIR
KLASIFIKASI PENYAKIT HIPERTENSI
PASIEN PUSKESMAS X KOTO I KABUPATEN TANAH DATAR
MENGGUNAKAN METODE *DECISION TREE* ALGORITMA C4.5

Nama : Abid Makarim Dyfa
NIM/Tahun Masuk : 19037002/2019
Program Studi : DIII Statistika
Fakultas : Matematika dan Ilmu Pengetahuan Alam

Padang, 5 Oktober 2022

Disetujui oleh :

Pembimbing Akademik



Dr. Syafriandi, M.Si.

NIP. 196609081991031003

HALAMAN PENGESAHAN LULUS UJIAN TUGAS AKHIR

Nama : Abid Makarim Dyfa
TM/NIM : 2019/19037002
Program Studi : DIII Statistika
Departemen : Statistika
Fakultas : Matematika dan Ilmu Pengetahuan Alam

**KLASIFIKASI PENYAKIT HIPERTENSI
PASIEN PUSKESMAS X KOTO I KABUPATEN TANAH DATAR
MENGUNAKAN METODE *DECISION TREE* ALGORITMA C4.5**

**Dinyatakan lulus setelah dipertahankan di depan Tim Penguji Tugas Akhir
Program Studi DIII Statistika Departemen Statistika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Negeri Padang**

Padang, 5 Oktober 2022

Tim Penguji

Nama

1. Ketua : Dr. Syafriandi, M.Si.
2. Anggota : Dr. Dony Permana, M.Si.
3. Anggota : Dra. Nonong Amalita, M.Si.

1. 
2. 
3. 

SURAT PERNYATAAN TIDAK PLAGIAT

Saya yang bertanda tangan di bawah ini :

Nama : Abid Makarim Dyfa
NIM/TM : 19037002/2019
Program Studi : DIII Statistika
Departemen : Statistika
Fakultas : MIPA UNP

Dengan ini menyatakan, bahwa Tugas Akhir saya dengan judul "**Klasifikasi Penyakit Hipertensi Pasien Puskesmas X Koto I Kabupaten Tanah Datar Menggunakan Metode *Decision Tree* Algoritma C4.5**" adalah benar merupakan hasil karya saya dan bukan merupakan plagiat dari karya orang lain atau pengutipan dengan cara-cara yang tidak sesuai dengan etika yang berlaku dalam tradisi keilmuan. Apabila suatu saat terbukti saya melakukan plagiat maka saya bersedia diproses dan menerima sanksi akademis maupun hukum sesuai dengan hukum dan ketentuan yang berlaku, baik di institusi UNP maupun di masyarakat dan Negara.

Demikianlah pernyataan ini saya buat dengan penuh kesadaran rasa tanggung jawab sebagai anggota masyarakat ilmiah.

Diketahui oleh,

Kepala Departemen Statistika,



Dr. Dony Permana, M.Si

NIP. 197501272006041001

Saya yang menyatakan,



Abid Makarim Dyfa

NIM. 19037002

**KLASIFIKASI PENYAKIT HIPERTENSI
PASIEN PUSKESMAS X KOTO I KABUPATEN TANAH DATAR
MENGUNAKAN METODE *DECISION TREE* ALGORITMA C4.5**

ABSTRAK

Abid Makarim Dyfa

Hipertensi adalah kondisi ketika tekanan darah pada pembuluh darah mengalir melebihi batas normal. Hipertensi adalah salah satu penyakit tidak menular yang merupakan *silent killer* atau pembunuh diam-diam, dikarenakan umumnya penderita hipertensi tidak menyadari sedang mengidap hipertensi. Berdasarkan Riskesdas provinsi Sumatera Barat 2018, prevalensi penderita hipertensi di Sumatera Barat sebesar 25,1%. Untuk prevalensi penderita hipertensi di Kabupaten Tanah datar usia 18 tahun ke atas sebesar 31,57%, yang mana angka ini melebihi angka prevalensi hipertensi Provinsi Sumatera Barat. Berdasarkan data Dinas Kesehatan Kabupaten Tanah Datar (2018), penyakit hipertensi menjadi salah satu penyebab kematian pada ibu hamil di tahun 2018. Hal yang mesti dilakukan adalah mengetahui ciri-ciri penderita hipertensi berdasarkan faktor penyebabnya.

Penelitian ini merupakan penelitian terapan, yaitu penelitian yang dilakukan berdasarkan kenyataan, penerapan dan pengembangan ilmu pengetahuan dasar dalam kehidupan nyata. Data yang digunakan pada penelitian yaitu data sekunder. Data yang dipakai adalah rekapitulasi data diagnosa penyakit hipertensi berdasarkan penyebab hipertensi dari pasien hipertensi Puskesmas X Koto I Kabupaten Tanah Datar. Variabel yang digunakan pada penelitian ini adalah Usia (X1), Jenis Kelamin (X2), Riwayat Keluarga penderita hipertensi (X3), Konsumsi Rokok (X4), Indeks Massa Tubuh (X5), dan Diagnosa hipertensi (Y). Langkah yang dilakukan yaitu memisah data latih dan data uji sebanyak 75% dan 25%. Pembentukan model pohon keputusan dengan nilai *entropy* dan *information gain*. Uji validasi model dengan *confusion matrix*.

Hasil penelitian didapat bahwa atribut usia yang menjadi penciri utama. Semua atribut yang dipakai terbukti mempengaruhi variabel target. Didapat nilai akurasi sebesar 0.71, artinya kualitas klasifikasi yang dilakukan berada pada tingkat “sedang”.

Kata Kunci: Hipertensi, Faktor Penyebab, Pohon Keputusan

**CLASSIFICATION OF HYPERTENSION DISEASES
PATIENTS OF PUSKESMAS X KOTO I DISTRICT TANAH DATAR
USING THE DECISION TREE ALGORITHM C4.5 METHOD**

ABSTRACT

Abid Makarim Dyfa

Hypertension is a condition when the blood pressure in the blood vessels exceeds the normal limit. Hypertension is a non-communicable disease which is a *silent killer*, because generally people with hypertension are not aware that they have hypertension. Based on Riskesdas of West Sumatra province 2018, the prevalence of hypertension sufferers in West Sumatra is 25.1%. The prevalence of hypertension sufferers in Tanah Datar Regency aged 18 years and over is 31.57%, which is higher than the prevalence rate of hypertension in West Sumatra Province. Based on data from the Tanah Datar District Health Office (2018), hypertension is one of the causes of death in pregnant women in 2018. What must be done is to know the characteristics of hypertension sufferers based on the causative factors.

This research is applied research, namely research conducted based on reality, application and development of basic science in real life. The data used in this research is secondary data. The data used is a recapitulation of data on the diagnosis of hypertension based on the cause of hypertension from hypertension patients at Puskesmas X Koto I, Tanah Datar Regency. The variables used in this study were Age (X1), Gender (X2), Family History of Hypertension Patients (X3), Cigarette Consumption (X4), Body Mass Index (X5), and Diagnosis of Hypertension (Y). The steps taken are to separate training data and test data by 75% and 25%. Formation of decision tree model with *entropy* and *information gain*. Model validation test with *confusion matrix*.

The results of the study found that the age attribute was the main characteristic. All attributes used are proven to affect the target variable. An accuracy value of 0.71 was obtained, meaning that the quality of the classification carried out was at the "fair" level.

Keywords: Hypertension, Causative Factor, Decision Tree

KATA PENGANTAR

Alhamdulillah puji syukur penulis panjatkan kepada Allah SWT Tuhan semesta alam yang telah melimpahkan rahmat, hidayat-Nya kepada penulis, sehingga penulis dapat menyelesaikan Tugas Akhir dengan judul “**Klasifikasi Penyakit Hipertensi Pasien Puskesmas X Koto I Menggunakan Metode *Decision Tree* Algoritma C4.5**” dengan baik. Serta Shalawat beserta salam penulis semoga tersampaikan kepada baginda Nabi Muhammad SAW.

Penulisan Tugas Akhir ini bertujuan untuk memenuhi salah satu syarat menyelesaikan pendidikan pada program studi Diploma III Statistika Departemen Statistika FMIPA UNP. Dalam pembuatan Tugas Akhir ini tidak terlepas dari bantuan dan dukungan dari berbagai pihak. Untuk itu, penulis mengucapkan rasa terima kasih yang sebesar-besarnya kepada yang terhormat :

1. Bapak Dr. Syafriandi, M.Si, dosen pembimbing dan dosen penasehat akademik yang telah memberikan arahan bimbingan ,dan nasehat kepada penulis dalam menyelesaikan Tugas Akhir.
2. Bapak Dr. Dony Permana, M.Si., Kepala Departemen Statistika dan dosen pembahas yang telah memberikan masukan dan arahan demi kesempurnaan Tugas Akhir ini.
3. Ibu Dra. Nonong Amalita, M.Si., Sekretaris Departemen Statistika dan dosen pembahas yang telah memberikan masukan dan arahan demi kesempurnaan Tugas Akhir ini.

4. Bapak Dodi Vionanda, M.Si, Ph.D., Koordinator Program Studi DIII Statistika Departemen Statistika FMIPA UNP.
5. Bapak dan Ibu Dosen, Staf Pengajar dan Karyawan Departemen Statistika FMIPA UNP yang telah membimbing dan berbagi ilmu pengetahuan kepada penulis selama perkuliahan.
6. Pimpinan dan Staf Puskesmas X Koto I yang telah mengizinkan penulis melakukan penelitian
7. Kedua orang tua penulis yang telah mendukung dan membantu penulis dalam menyelesaikan Tugas Akhir baik dukungan moril dan dukungan yang lainnya
8. Teman, dan rekan-rekan yang telah memberikan dukungan kepada penulis.

Semoga bantuan dan bimbingan yang diberikan kepada penulis menjadi amal ibadah dan dibaloh oleh Allah SWT. Penulis menyadari bahwa penulisan Tugas Akhir ini jauh dari kesempurnaan. Oleh karena itu, saran dan masukan dari berbagai pihak sangat diharapkan demi kesempurnaan penulisan yang akan datang. Akhir kata penulis berharap Tugas Akhir ini dapat memberikan sumbangan pikiran kepada pembaca.

Padang, Oktober 2022

Abid Makarim Dyfa

DAFTAR ISI

ABSTRAK	i
KATA PENGANTAR.....	iii
DAFTAR ISI.....	v
DAFTAR TABEL.....	vii
DAFTAR GAMBAR.....	viii
DAFTAR LAMPIRAN	ix
BAB I PENDAHULUAN	1
A. Latar Belakang Masalah	1
B. Rumusan Masalah.....	5
C. Tujuan Penelitian.....	5
D. Batasan Masalah	5
E. Manfaat Penelitian.....	6
BAB II KERANGKA TEORETIS	7
A. Landasan Teori	7
1. Pengertian Hipertensi	7
2. Pengelompokan Tingkatan Hipertensi.....	8
3. Bahaya Hipertensi.....	9
4. Penyebab Hipertensi	10
5. Data Mining	13
6. Klasifikasi	16
7. Pohon Keputusan (Decision Tree).....	18
8. Algoritma C4.5	19
9. Validasi	23
B. Penelitian Relevan	26
BAB III METODOLOGI PENELITIAN	28
A. Jenis Penelitian	28
B. Jenis dan Sumber Data	28
C. Populasi dan Sampel.....	28
1. Populasi.....	28
2. Sampel	28

D. Variabel Penelitian	29
E. Struktur Data.....	30
F. Teknik Analisis Data	30
BAB IV HASIL DAN PEMBAHASAN.....	32
A. Hasil.....	32
1. Statistika Deskriptif.....	32
2. Analisis Data.....	36
B. Pembahasan	53
BAB V PENUTUP	55
A. Kesimpulan.....	55
B. Saran	56
DAFTAR PUSTAKA	57
LAMPIRAN.....	59

DAFTAR TABEL

Tabel	Halaman
1. Kelompok Besaran Tekanan Darah Menurut WHO.....	9
2. Confusion Matrix.....	22
3. Variabel, Keterangan dan Tipe Data.....	30
4. Struktur Data.....	30
5. Ukuran Pemusatan dan Penyebaran Data.....	36
6. Rangkuman Perhitungan Entropy dan Information Gain Node Akar.....	42
7. Rangkuman Perhitungan Entropy dan Information Gain Usia Lebih dari 50 Tahun.....	47
8. Rangkuman Perhitungan Entropy dan Information Gain Usia Kurang atau 50 Tahun.....	51
9. Hasil Confusion Matrix.....	52
10. Aturan Klasifikasi Diagnosa Hipertensi.....	53

DAFTAR GAMBAR

Gambar	Halaman
1. Prevalensi Penderita Hipertensi di Indonesia Tahun 2013 dan Tahun 2018.....	1
2. Prevalensi Penderita Hipertensi di Sumatera Barat dan Kabupaten Tanah Datar.....	2
3. Alur Kerja Data Mining.....	14
4. Model Pohon Keputusan/ Decision Tree.....	18
5. Jumlah Pasien Hipertensi Berdasarkan Riwayat Keluarga Penderita Hipertensi dan Diagnosa Hipertensi.....	32
6. Jumlah Pasien Hipertensi Berdasarkan Jenis Kelamin dan Diagnosa Hipertensi.....	33
7. Jumlah Pasien Hipertensi Berdasarkan Konsumsi Rokok dan Diagnosa Hipertensi.....	34
8. Jumlah Pasien Hipertensi Berdasarkan Kriteria Indeks Massa Tubuh dan Diagnosa Hipertensi.....	39
9. Jumlah Pasien Hipertensi Berdasarkan Usia dan Diagnosa Hipertensi.....	39
10. Pohon Keputusan Klasifikasi Penyakit Hipertensi	52

DAFTAR LAMPIRAN

Lampiran	Halaman
1. Data Penelitian.....	59
2. Output Pohon Keputusan Menggunakan Aplikasi Rapidminer.....	60
3. Output Confusion Matrix Menggunakan Aplikasi Rapidminer.....	61

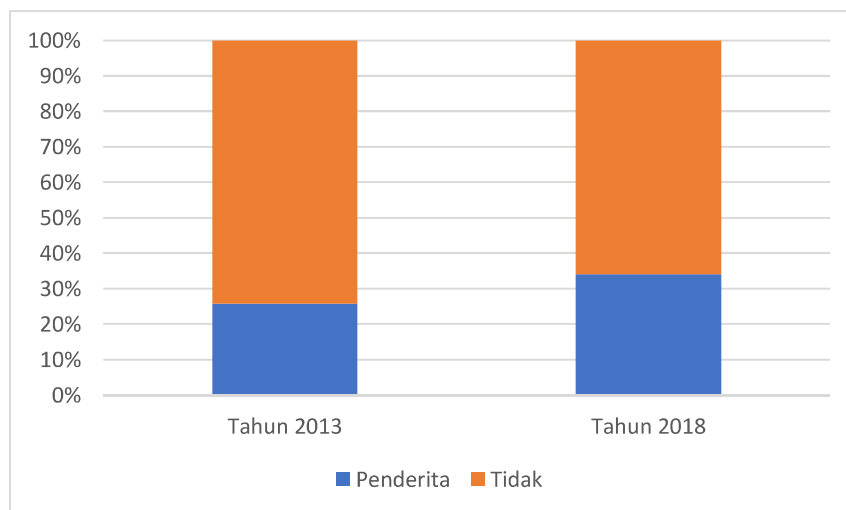
BAB I

PENDAHULUAN

A. Latar Belakang Masalah

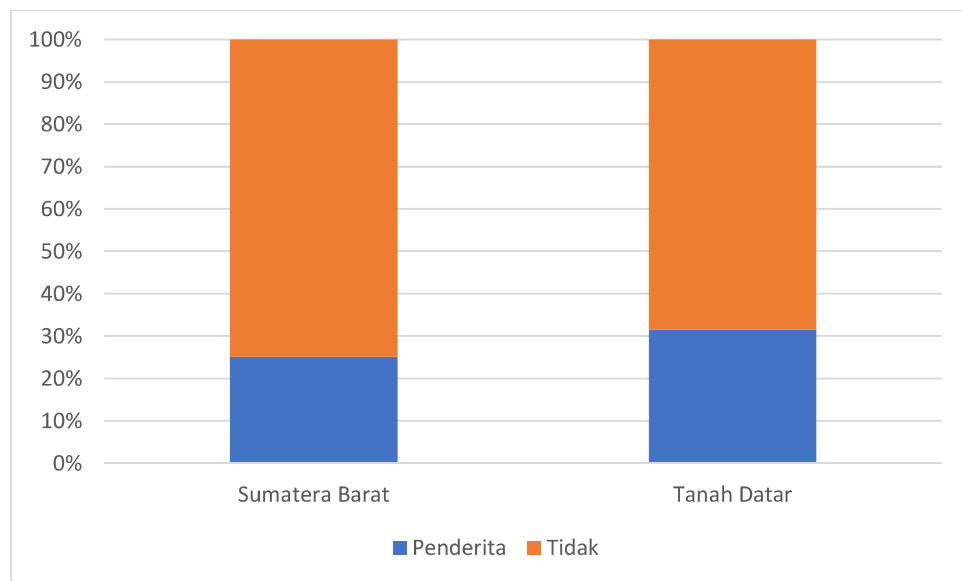
Hipertensi adalah kondisi ketika tekanan darah pada pembuluh darah mengalir melebihi batas normal (Papadakis: 2019, 451). Besaran tekanan darah normal yaitu 120/80 mmHg. Hipertensi adalah salah satu penyakit tidak menular yang merupakan *silent killer* atau pembunuh diam-diam menurut *World Health Organization* (WHO). Hal ini dikarenakan umumnya penderita hipertensi tidak menyadari sedang mengidap hipertensi.

Berdasarkan hasil Riset Kesehatan Dasar Nasional (Riskesdas) 2018, angka prevalensi penderita hipertensi di Indonesia mencapai 34%. Angka ini naik dari hasil Riskesdas sebelumnya di tahun 2013, yang mana angka prevalensi penderita hipertensi pada tahun 2013 sebesar 25,8%. Berikut tampilan grafik prevalensi penderita hipertensi di Indonesia pada Gambar 1.



Gambar 1. Prevalensi Penderita Hipertensi di Indonesia Tahun 2013 dan Tahun 2018

Hasil Riskesdas provinsi Sumatera Barat 2018, menunjukkan angka prevalensi penderita hipertensi di Sumatera Barat sebesar 25,1%. Angka prevalensi penderita hipertensi di Kabupaten Tanah datar usia 18 tahun ke atas sebesar 31,57%, yang mana angka ini melebihi angka prevalensi penderita hipertensi Provinsi Sumatera Barat. Gambar 2 menampilkan grafik prevalensi penderita hipertensi di Provinsi Sumatera Barat dan Kabupaten Tanah Datar.



Gambar 2. Prevalensi Penderita Hipertensi di Provinsi Sumatera Barat dan Kabupaten Tanah Datar

Berdasarkan data Dinas Kesehatan Kabupaten Tanah Datar (2018), penyakit hipertensi menjadi salah satu penyebab kematian pada ibu hamil di tahun 2018. Berdasarkan hasil rekam medis pada Puskesmas X Koyo I tahun 2021, hipertensi merupakan penyakit tidak menular dengan jumlah penderita terbanyak ketiga setelah diabetes dan penyakit jantung, yaitu dengan jumlah 930 orang.

Tingginya angka penderita hipertensi tidak lepas dari karakteristiknya yaitu *silent killer* dengan banyaknya penderita yang tidak menyadari sedang mengidap hipertensi, dan menganggap remeh hipertensi.

Gejala yang dirasakan pada penderita hipertensi jarang timbul pada awalnya. Menurut Susilo (2011, 23), ketika hipertensi sudah pada tahap stadium 2, biasanya akan timbul gejala seperti kepala pusing, jantung berdebar, dan terasa nyeri di bagian belakang kepala. Mayoritas dari penderita hipertensi stadium 1 tidak merasakan sama sekali gejala hipertensi. Hal inilah yang membuat lambat laun hipertensi yang diderita semakin parah dan berpotensi terkena gangguan kronis.

Penyakit hipertensi akan lebih baik dicegah sejak awal untuk memperkecil kemungkinan terserang penyakit kronis. Penyakit hipertensi memang tidak menyebabkan kematian secara langsung, namun hipertensi menjadi penunjang terbesar komplikasi penyakit serius (Hull: 1993, 20). Adapun penyakit serius yang dipicu oleh hipertensi yaitu penyakit jantung, gagal ginjal, stroke, pembuluh darah pecah dan lainnya. Penyakit tersebutlah yang berperan langsung menyebabkan kematian.

Hipertensi bisa disebabkan oleh faktor internal dan faktor eksternal seseorang. Menurut Susuilo (2011,48), faktor internal adalah faktor yang berasal dari dalam tubuh seseorang, dan tidak dapat dikendalikan, sedangkan faktor eksternal adalah faktor yang berasal dari luar tubuh seseorang. Contoh faktor internal yang memicu hipertensi yaitu usia, jenis kelamin, dan genetik. Beberapa faktor eksternal yang menjadi pemicu hipertensi yaitu konsumsi rokok dan kegemukan.

Hipertensi memang tergolong penyakit tidak menular, namun jumlah penderita hipertensi tidak bisa diremehkan. Untuk mengantisipasi angka penderita hipertensi yang semakin meningkat, hal yang mesti dilakukan yaitu

mengetahui ciri-ciri penderita hipertensi berdasarkan faktor internal dan faktor eksternal dengan mengelompokkan tingkatan hipertensi yang telah diklasifikasikan oleh WHO. Cara yang digunakan adalah klasifikasi.

Hal pertama yang harus diketahui untuk mempelajari dasar klasifikasi yaitu mempelajari *data mining*. Data mining adalah proses pemanfaatan data dalam basis data dengan mengolahnya sehingga menghasilkan informasi baru yang berguna. Menurut Prasetyo (2013,11), klasifikasi adalah mengelompokkan beberapa atribut menjadi beberapa kelas oleh klasifikator dengan aturan yang telah ditetapkan.

Ada banyak metode dalam klasifikasi, diantaranya yaitu *Decision Tree*, *Naïve Bayes*, *Support Vector Machine*, dan *Random Forest*. Metode *Naïve Bayes*, *Support Vector Machine*, dan *Random Forest*. Metode *Decision Tree* atau Pohon Keputusan menampilkan model seperti pohon terbalik dengan akar terletak paling atas. Akar melambangkan atribut yang paling signifikan dalam mempengaruhi variabel target, sehingga waktu yang dibutuhkan untuk mengetahui penciri utama akan lebih singkat.

Ada beberapa metode dalam membangun *decision tree*, yaitu CART (*Classification and Regression Tree*), ID3 (*Iterative Dichotomiser 3*), dan Algoritma C4.5. Metode CART adalah metode *decision tree* yang melakukan percabangan biner. Hasil dari model pohon keputusan menggunakan metode CART selalu membagi cabang menjadi 2. Ketika ada atribut yang memiliki cabang lebih dari 2, maka perlu dilakukan pemilihan cabang terbaik sehingga akan memakan waktu pengerjaan menjadi lebih lama. Pada ID3, induksi *decision tree* hanya bisa mengatasi data dengan tipe kategorikal (nominal atau

ordinal), belum mampu digunakan pada tipe data numerikal (interval atau rasio). Algoritma C4.5 pertama kali dikenalkan oleh Quinlan (1996) sebagai versi perbaikan dari ID3 yakni sudah bisa mengatasi data dengan tipe numerik. Algoritma C4.5 berbeda dengan CART dari segi cara menurunkan cabangnya, pada Algoritma C4.5 menurunkan cabang dengan semua kategori atribut pada cabang.

B. Rumusan Masalah

Berdasarkan uraian latar belakang masalah, maka rumusan masalah yang terbentuk adalah:

1. Apa saja penciri penderita hipertensi berdasarkan klasifikasi yang dilakukan?
2. Apa penciri utama penderita hipertensi pasien Puskesmas X Koto I?

C. Tujuan Penelitian

Tujuan dari penelitian ini adalah:

1. Mengetahui karakteristik pengidap hipertensi pada masing masing tingkatan hipertensi.
2. Mengetahui karakteristik utama penderita hipertensi pasien Puskesmas X Koto I

D. Batasan Masalah

Batasan masalah pada penelitian ini yaitu:

1. Metode klasifikasi yang digunakan adalah metode *decision tree* menggunakan algoritma C4.5
2. Klasifikasi yang dilakukan menggunakan data pasien hipertensi rawat jalan Puskesmas X Koto I.

E. Manfaat Penelitian

Manfaat penelitian ini adalah:

1. Bagi Penulis

Penulisan ini ditujukan sebagai salah satu syarat lulus pendidikan Diploma III Departemen Statistika Universitas Negeri Padang. Tugas akhir ini sebagai penerapan ilmu yang telah dipelajari selama masa perkuliahan.

2. Bagi Pembaca

Diharapkan Tugas Akhir ini bisa dijadikan tambahan referensi dan bahan acuan yang dapat dikembangkan lagi bagi calon peneliti selanjutnya dalam menerapkan metode *decision tree* dengan algoritma C4.5.

3. Bagi Puskesmas X Koto I Kabupaten Tanah Datar

Sebagai hasil akhir dari tugas akhir ini berupa klasifikasi tingkatan hipertensi, semoga bisa membantu dan bahan pendukung pihak Puskesmas X Koto I dalam mengklasifikasi pengidap hipertensi.