

**KLASIFIKASI FAKTOR PENYEBAB PENYAKIT HIPERTENSI  
DI PUSKESMAS GUGUK PANJANG KOTA BUKITTINGGI  
MENGUNAKAN *CLASSIFICATION AND REGRESSION TREE***

**TUGAS AKHIR**

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



**Oleh**

**ADINDA NAFIZAH CINTA UGRA  
NIM 20037003**

**PROGRAM STUDI DIPLOMA III STATISTIKA  
DEPARTEMEN STATISTIKA  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS NEGERI PADANG  
2024**

**PERSETUJUAN TUGAS AKHIR**

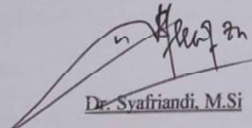
**KLASIFIKASI FAKTOR PENYEBAB HIPERTENSI DI  
PUSKESMAS GUGUK PANJANG KOTA BUKITTINGGI  
MENGUNAKAN *CLASSIFICATION AND REGRESSION TREE***

Nama : Adinda Nafizah Cinta Ugra  
NIM : 20037003  
Program Studi : D3 Statistika  
Departemen : Statistika  
Fakultas : Matematika dan Ilmu Pengetahuan Alam

Padang, 26 Februari 2024

Disetujui Oleh:

Pembimbing



Dr. Syafriandi, M.Si

NIP. 19660908 199103 1 003

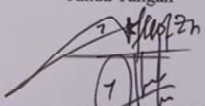
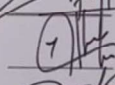
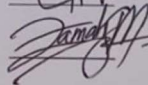
**PENGESAHAN LULUS UJIAN TUGAS AKHIR**

Nama : Adinda Nafizah Cinta Ugra  
NIM : 20037003  
Program Studi : D3 Statistika  
Departemen : Statistika  
Fakultas : Matematika dan Ilmu Pengetahuan Alam

**KLASIFIKASI FAKTOR PENYEBAB HIPERTENSI DI  
PUSKESMAS GUGUK PANJANG KOTA BUKITTINGGI  
MENGUNAKAN *CLASSIFICATION AND REGRESSION TREE***

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

Padang, 26 Februari 2024

	Tim Penguji	Tanda Tangan
	Nama	
Ketua	: Dr. Syafriandi, M.Si	
Anggota	: Dr. Yenni Kurniawati, M.Si	
Anggota	: Zamahsary Martha, S.Si., M.Si	

## SURAT PERNYATAAN TIDAK PLAGIAT

Saya yang bertandatangan di bawah ini:

Nama : Adinda Nafizah Cinta Ugra  
NIM : 20037003  
Program Studi : D3 Statistika  
Departemen : Statistika  
Fakultas : Matematika dan Ilmu Pengetahuan Alam

Dengan ini menyatakan bahwa, skripsi saya dengan judul "**Klasifikasi Faktor Penyebab Penyakit Hipertensi di Puskesmas Guguk Panjang Kota Bukittinggi Menggunakan Metode *Classification And Regression Tree***" 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.

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

Diketahui oleh,  
Kepala Departemen Statistika,



Dr. Yenni Kurniawati, S.Si., M.Si  
NIP. 1984022320101220005

Saya yang menyatakan,


Adinda Nafizah Cinta Ugra  
NIM. 20037003

# Klasifikasi Faktor Penyebab Penyakit Hipertensi di Puskesmas Guguk Panjang Kota Bukittinggi Menggunakan *Classification And Regression Tree*

Adinda Nafizah Cinta Ugra

## ABSTRAK

Hipertensi adalah salah satu penyakit tidak menular yang merupakan *silent killer* atau pembunuh diam-diam. *World Health Organization* (WHO) melaporkan bahwa 9,4 juta orang meninggal setiap tahunnya akibat hipertensi. Penderita hipertensi umumnya tidak menyadari sedang mengidap hipertensi. Faktor penyebab seseorang terkena hipertensi dipengaruhi dua faktor yaitu, faktor internal dan faktor eksternal. Faktor internal terdiri dari umur, jenis kelamin, sedangkan faktor eksternal terdiri dari indeks masa tubuh, tekanan sistolik dan tekanan diastolik.

Penelitian ini merupakan penelitian terapan. Data yang digunakan adalah data sekunder. Data yang dipakai adalah rekapitulasi data pasien kontrol rutin di Puskesmas Guguk Panjang Kota Bukittinggi pada bulan Januari-Juli 2023. Variabel yang digunakan pada penelitian ini adalah Umur ( $X_1$ ), Jenis Kelamin ( $X_2$ ), Indeks Massa Tubuh ( $X_3$ ), Tekanan Sistolik ( $X_4$ ), Tekanan Diastolik ( $X_5$ ), dan Kriteria Penyakit (Y).

Hasil penelitian menggunakan metode CART menghasilkan pohon keputusan optimal dengan algoritma CART menghasilkan sepuluh keputusan dimana terdapat lima keputusan yang hipertensi dan lima keputusan yang tidak hipertensi. Faktor yang berpengaruh pada penelitian ini yaitu variabel Indeks Masa Tubuh dengan keakuratan model sebesar 88,75 %.

**Kata Kunci** : Hipertensi, Klasifikasi, Algoritma CART

# **Classification of Causal Factors of Hypertension in Puskesmas Guguk Panjang Bukittinggi City Using Classification And Regression Tree**

Adinda Nafizah Cinta Ugra

## **ABSTRACT**

Hypertension is a non-communicable disease that is a silent killer. The World Health Organization (WHO) reports that 9.4 million people die each year from hypertension. Hypertension sufferers generally do not realize they have hypertension. The factors that cause a person to develop hypertension are influenced by two factors, namely, internal factors and external factors. Internal factors consist of age, gender, while external factors consist of body mass index, systolic pressure and diastolic pressure.

This research is applied research. The data used is secondary data. The data used is a recapitulation of routine control patient data at the Guguk Panjang Health Center, Bukittinggi City in January-July 2023. The variables used in this study are Age ( $X_1$ ), Gender ( $X_2$ ), Body Mass Index ( $X_3$ ), Systolic Pressure ( $X_4$ ), Diastolic Pressure ( $X_5$ ), and Disease Criteria ( $Y$ ).

The results of research using the CART method produce an optimal decision tree with the CART algorithm producing ten decisions where there are five hypertensive decisions and five decisions that are not hypertensive. The influential factor in this study is the Body Mass Index variable with a model accuracy of 88.75%.

**Keywords:** Hypertension, Classification, CART Algorithm

## KATA PENGANTAR

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Puji dan syukur penulis ucapkan kehadiran Allah SWT yang telah memberikan rahmat dan hidayah-Nya serta kesehatan kepada penulis untuk melaksanakan Tugas Akhir ini dengan judul: **“Klasifikasi Faktor Penyebab Hipertensi di Puskesmas Guguk Panjang Kota Bukittinggi Menggunakan *Classification And Regression Tree*”** dapat terselesaikan dengan baik.

Tugas Akhir ini dibuat sebagai salah satu penyelesaian dari mata kuliah Program Studi Diploma III Statistika Departemen Statistika FMIPA Universitas Negeri Padang. Selama penyelesaian Tugas Akhir penulis telah banyak mendapatkan bantuan pengarahan, saran dan motivasi dari berbagai pihak, untuk itu penulis mengucapkan terima kasih banyak kepada:

1. Bapak Dr. Syafriandi, M.Si., sebagai penasehat akademik yang telah membimbing, memberi saran, dan nasehat dalam menyelesaikan Tugas Akhir ini.
2. Ibu Dr. Yenni Kurniawati, M.Si., dosen pembahas yang telah memberikan masukan dan arahan dalam pengerjaan Tugas Akhir ini, sekaligus Kepala Departemen Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam.
3. Bapak Zamahsary Martha, M.Si., dosen pembahas yang telah memberikan masukan dan arahan dalam pengerjaan Tugas Akhir ini.
4. Bapak Dodi Vionanda, M.Si., Ph.D., Koordinator Program Studi Diploma III Statistika Departemen Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam.

5. Bapak, ibu dosen, dan tenaga kependidikan Departemen Statistika yang telah membimbing dan berbagai ilmu pengetahuan kepada penulis selama duduk di bangku perkuliahan.
6. Terutama sekali kepada kedua orang tua dan saudara-saudara tercinta yang telah memberikan dukungan baik secara moril dan material serta doa restu kepada penulis selama duduk di bangku perkuliahan sehingga penulis dapat menyelesaikan Tugas Akhir ini.
7. Terima kasih kepada sahabat terdekat dan teman-teman Diploma III Statistika 2020 yang telah membantu dan memberikan semangat dan saran dalam menyelesaikan Tugas Akhir ini.

Penulis menyadari bahwa penulisan serta penyusunan Tugas Akhir ini masih jauh dari kata sempurna. Karena itu, penulis mohon masukan, kritik dan saran agar menjadi bahan introspeksi diri kemudian hari. Semoga Tugas Akhir ini bermanfaat bagi penulis dan pembaca pada umumnya.

Padang, 26 Februari 2024

Penulis

Adinda Nafizah Cinta Ugra



## DAFTAR ISI

<b>ABSTRAK .....</b>	<b>i</b>
<b>KATA PENGANTAR.....</b>	<b>iii</b>
<b>DAFTAR ISI.....</b>	<b>v</b>
<b>DAFTAR TABEL .....</b>	<b>vii</b>
<b>DAFTAR GAMBAR.....</b>	<b>viii</b>
<b>DAFTAR LAMPIRAN.....</b>	<b>ix</b>
<b>BAB I PENDAHULUAN.....</b>	<b>1</b>
A. Latar Belakang Masalah.....	1
B. Batasan Masalah.....	6
C. Rumusan Masalah .....	6
D. Tujuan Penelitian .....	6
E. Manfaat Penelitian .....	7
<b>BAB II KERANGKA TEORITIS .....</b>	<b>8</b>
A. Hipertensi .....	8
B. Faktor - Faktor Penyebab Hipertensi .....	10
C. Data Mining .....	12
D. Klasifikasi .....	13
E. Pohon Keputusan .....	14
F. <i>Preprocessing</i> Data .....	14
G. <i>Metode Classification And Regression Tree</i> .....	16
H. Bentuk dan Struktur Pohon Klasifikasi.....	23
I. <i>Confusion Matrix</i> .....	23
<b>BAB III METODOLOGI PENELITIAN .....</b>	<b>26</b>
A. Jenis Penelitian.....	26
B. Jenis dan Sumber Data .....	26
C. Prosedur Penelitian.....	26
D. Teknik Analisis Data.....	27
<b>BAB IV HASIL DAN PEMBAHASAN .....</b>	<b>29</b>
A. Hasil Penelitian .....	29
B. Pembahasan.....	46

<b>BAB V PENUTUP.....</b>	<b>50</b>
A. Kesimpulan .....	50
B. Saran.....	51
<b>DAFTAR PUSTAKA .....</b>	<b>52</b>

## DAFTAR TABEL

<b>TABEL</b>	<b>Halaman</b>
1. Kelompok Besaran Tekanan Darah Menurut WHO.....	9
2. <i>Confussion Matrix</i> .....	24
3. Variabel Penelitian .....	27
4. Struktur Data.....	27
5. Calon Simpul Pemilah.....	34
6. Calon Simpul Variabel Indeks Masa Tubuh.....	38
7. <i>Goodness of Split</i> .....	39
8. <i>Complexcity Parameter Simpul internal Pohon Maksimal</i> .....	42
9. Nilai Complexity Parameter Pohon Keputusan CART .....	43
10. Uji Ketepatan Klasifikasi .....	46

## DAFTAR GAMBAR

GAMBAR	Halaman
1. Prevalensi Penderita Hipertensi di Indonesia pada Tahun 2013 dan Tahun 2018 .....	1
2. Prevalensi Penderita Hipertensi di Sumatera Barat dan Kota Bukittinggi pada Tahun 2018 .....	2
3. Struktur Pohon Klasifikasi CART .....	23
4. Deskripsi Persentase Penyakit Hipertensi Pada Bulan Januari-Juni 2023	30
5. Klasifikasi Kriteria Penyakit Berdasarkan Umur Pasien.....	30
6. Klasifikasi Kriteria Penyakit Pasien Berdasarkan Jenis Kelamin .....	31
7. Klasifikasi Kriteria Penyakit Pasien Berdasarkan Indeks Masa Tubuh ....	31
8. Klasifikasi Kriteria Penyakit Pasien Berdasarkan Tekanan Sistolik .....	32
9. Klasifikasi Kriteria Penyakit Pasien Berdasarkan Tekanan Diastolik .....	32
10. Pohon CART Pembelahan Pertama.....	41
11. Pohon Keputusan CART Maksimal .....	41
12. <i>Plot Complexity parameter dan Cross Validation Relative Error</i> .....	43
13. Pohon Optimal.....	45

## DAFTAR LAMPIRAN

LAMPIRAN	Halaman
1. Data Penelitian.....	54
2. Proses Splitting data latih dan data uji menggunakan R studio.....	55
3. Proses pembangunan model pohon maksimal menggunakan R studio ..	57
4. Visualisasi Pohon Keputusan CART maksimal .....	59
5. Proses Pruning Pohon Maksimal .....	60
6. Visualisasi Pohon Pruning .....	61
7. Uji Ketepatan Pohon Optimal.....	62

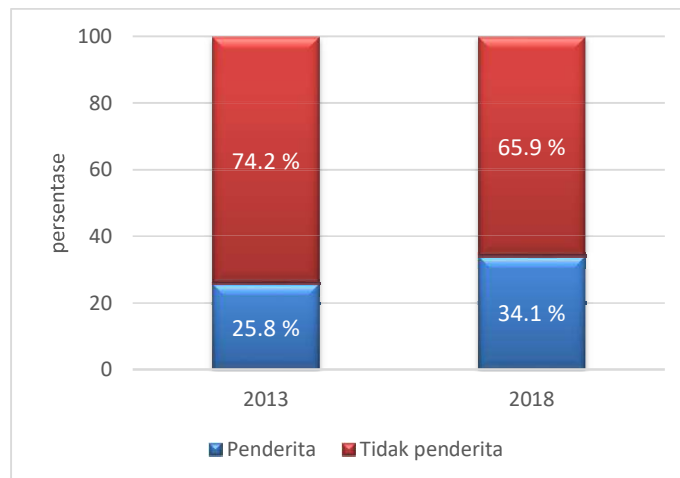
# BAB I

## PENDAHULUAN

### A. Latar Belakang Masalah

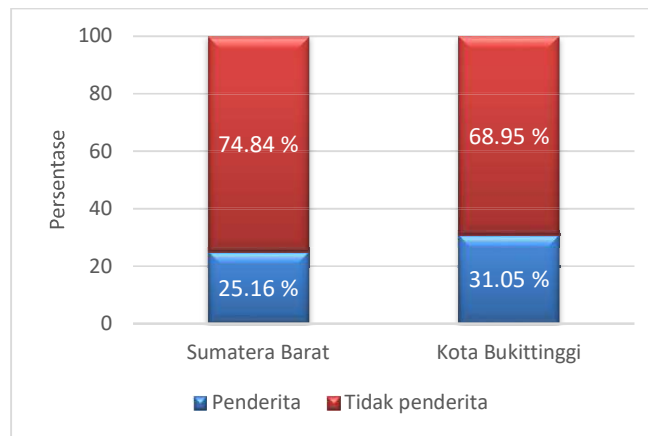
Hipertensi adalah kondisi ketika tekanan darah pada pembuluh darah mengalir melebihi batas normal (Papadakis, 2019:451). Tekanan darah normal yaitu 120/80 mmHg. Menurut *World Health Organization* (WHO) hipertensi adalah salah satu penyakit tidak menular yang merupakan *silent killer* atau pembunuh diam-diam. WHO melaporkan bahwa 9,4 juta orang meninggal setiap tahunnya akibat hipertensi. Hal ini dikarenakan umumnya penderita hipertensi tidak menyadari sedang mengidap hipertensi.

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



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

Berdasarkan hasil Riskedas tahun 2018, menunjukkan angka prevalensi penderita hipertensi di Provinsi Sumatera Barat sebesar 25,16%. Angka prevalensi penderita hipertensi di Kota Bukittinggi usia 18 tahun ke atas sebesar 31,05%, yang mana angka prevalensi penderita hipertensi di Kota Bukittinggi melebihi angka prevalensi penderita Provinsi Sumatera Barat. Gambar 2 menampilkan grafik prevalensi penderita hipertensi di Provinsi Sumatera Barat dan Kota Bukittinggi.



**Gambar 2.** Prevalensi Penderita Hipertensi di Sumatera Barat dan Kota Bukittinggi pada Tahun 2018

Berdasarkan laporan profil kesehatan Kota Bukittinggi (2020), beberapa penyakit tidak menular yang banyak di Kota Bukittinggi adalah infeksi saluran pernafasan akut (ISPA), hipertensi, diabetes melitus, kanker payudara, dan jantung koroner. Hipertensi merupakan penyakit tidak menular dengan jumlah penderita terbanyak kedua dengan jumlah 15.704 kasus. Berdasarkan rekap medis Puskesmas Guguk Panjang Kota Bukittinggi tahun 2022, hipertensi merupakan penyakit tidak menular dengan jumlah penderita terbanyak, yaitu dengan jumlah 1.733 orang.

Tingginya angka penderita hipertensi tidak lepas dari karakteristiknya yaitu *silent killer* banyaknya penderita yang tidak menyadari sedang mengidap hipertensi, dan menganggap remeh hipertensi. 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 penderita hipertensi stadium 1 tidak merasakan sama sekali gejala hipertensi. Hal inilah yang membuat perlahan hipertensi yang diderita semakin parah dan berpotensi terkena gangguan kronis.

Penyakit hipertensi 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 lebih baik dicegah sejak awal untuk memperkecil terjadinya penyakit kronis. Hipertensi bisa disebabkan oleh dua faktor yaitu faktor eksternal dan faktor internal. Menurut Susilo (2018:48) faktor eksternal adalah faktor yang berasal dari luar tubuh seseorang, sedangkan faktor internal adalah faktor yang berasal dari dalam tubuh seseorang, dan tidak dapat dikendalikan. Faktor eksternal yang menjadi pemicu penyakit hipertensi yaitu kegemukan, sedangkan faktor internal yang menjadi penunjang yaitu usia, jenis kelamin, dan genetik.

Hipertensi memang tergolong penyakit tidak menular, namun jumlah penderita hipertensi tidak bisa dianggap remeh. Untuk mengantisipasi angka penderita hipertensi yang semakin meningkat, hal yang mesti dilakukan mengetahui ciri-ciri penderita hipertensi berdasarkan faktor-faktornya. Salah satu analisis yang dapat dilakukan adalah klasifikasi. Klasifikasi merupakan sebuah teknik yang menghasilkan model atau fungsi yang dapat membedakan sekumpulan data dan menghasilkan sekumpulan data yang belum pernah dipelajari (Mulyanto, 2011).



Metode klasifikasi pada penelitian ini digunakan untuk membantu dalam memahami pengelompokan data pasien yang menderita penyakit hipertensi. Menurut penelitian Wang dan Liu (2019) faktor-faktor yang menyebabkan hipertensi adalah jenis kelamin, umur, riwayat keluarga, tekanan sistolik, tekanan diastolik, dan kadar gula darah. Hasil klasifikasi ini akan mengidentifikasi pola-pola yang signifikan dalam data penelitian ini menunjukkan akurasi yang baik dalam mengidentifikasi risiko hipertensi.

Algoritma klasifikasi yang dapat digunakan diantaranya *Support Vektor Machine*, *Bayesian Methode*, *Decission Tree*, *Fuzzy*, dan Jaringan Saraf Tiruan. Salah satu bentuk keputusan dari analisis klasifikasi yang dapat diinterpretasikan dengan baik adalah *Decission Tree* (pohon keputusan). Beberapa algoritma yang dipakai untuk menghasilkan pohon keputusan, adalah *Chi-Squared Automatic Interaction Detector* (CHAID), *Commercial Version 5.0* (C5.0), dan *Classification and Regression Tree* (CART). Menurut penelitian yang dilakukan oleh Pratiwi (2020), yang membandingkan antara algoritma C5.0 dan CART, didapatkan hasil bahwa akurasi algoritma CART lebih tinggi dari pada akurasi dari algoritma C5.0. Penelitian sebelumnya juga dilakukan oleh Aritongan dkk (2016) yang membandingkan antara CART dan CHAID, didapatkan hasil bahwa algoritma CART memiliki akurasi yang lebih tinggi dibandingkan dengan algoritma CHAID.

*Classification and Regression Tree* (CART) adalah salah satu metode atau algoritma dari teknik pohon keputusan (*Decision Tree*). CART adalah salah satu metode statistik nonparametrik yang dapat menggambarkan hubungan antara variabel dependen (respon) dengan satu atau lebih variabel independen (prediktor) yang meliputi data diskrit, dan kontinu. Menurut Breiman (1984), apabila variabel

respon berskala kontinu atau diskrit, maka metode yang digunakan adalah CART. Metode ini akan menghasilkan pohon klasifikasi apabila peubah responnya diskrit dan menghasilkan pohon regresi apabila peubah responnya kontinu. Tujuan utama CART adalah untuk mendapatkan suatu kelompok data yang akurat sebagai penciri dari suatu pengklasifikasian. CART memberikan kemudahan untuk menginterpretasikan hasil analisis dan memberikan dugaan dengan tingkat kesalahan yang kecil (Sartono dan Syafitri, 2010).

Beberapa penelitian yang relevan dengan penelitian ini adalah penelitian yang dilakukan oleh Subarkah, P. (2020) penerapan algoritma klasifikasi *classification and regression tree* (CART) untuk diagnosis penyakit diabetes *retinopathy*, Sumartini, S. H., dan Purnami, S. W. (2015) juga telah menggunakan metode CART untuk klasifikasi rekurensi pasien kanker serviks di RSUD Dr. Soetomo Surabaya. Oleh karena itu penelitian ini akan digunakan algoritma CART untuk mengklasifikasikan faktor-faktor yang menjadi penyebab hipertensi.

Berdasarkan permasalahan yang telah diuraikan sebelumnya, akan dilakukan penanganan masalah pada pengklasifikasian data penyakit hipertensi di Puskesmas Guguk Panjang. Oleh karena itu, penulis melakukan penelitian yang berjudul **“Klasifikasi Faktor-faktor Penyakit Hipertensi di Puskesmas Guguk Panjang Kota Bukittinggi Menggunakan Metode *Classification and Regression Tree*”**.

## **B. Batasan Masalah**

Batasan masalah pada penelitian ini adalah.

1. Metode klasifikasi yang digunakan adalah *decision tree* menggunakan metode CART.
2. Klasifikasi yang dilakukan menggunakan data pasien hipertensi bulan Januari-Juli 2023 di UPTD Puskesmas Guguk Panjang.

## **C. Rumusan Masalah**

Rumusan masalah dalam penelitian ini adalah sebagai berikut.

1. Bagaimana hasil klasifikasi faktor-faktor penyebab hipertensi di Puskesmas Guguk Panjang Kota Bukittinggi menggunakan metode *Classification and Regression Tree*.
2. Bagaimana tingkat akurasi klasifikasi penyakit hipertensi di Puskesmas Guguk Panjang Kota Bukittinggi dengan Metode *Classification and Regression Tree*.

## **D. Tujuan Penelitian**

Tujuan dari penelitian ini adalah sebagai berikut.

1. Untuk mengetahui hasil klasifikasi faktor-faktor penyebab hipertensi di Puskesmas Guguk Panjang Kota Bukittinggi menggunakan metode *Classification and Regression Tree*.
2. Untuk mengetahui tingkat akurasi klasifikasi penyakit hipertensi di Puskesmas Guguk Panjang Kota Bukittinggi dengan Metode *Classification and Regression Tree*.

## **E. Manfaat Penelitian**

Manfaat penelitian ini adalah sebagai berikut.

### 1. Bagi Penulis

Menambah wawasan penulis dalam memahami analisis *Classification and Regression Tree* sebagai penerapan ilmu yang telah dipelajari selama masa perkuliahan.

### 2. Bagi Pembaca

Sebagai tambahan referensi dan bahan acuan yang dapat dikembangkan bagi calon peneliti selanjutnya dalam menerapkan *decision tree* metode *Classification and Regression Tree*.

### 3. Bagi Puskesmas Guguk Panjang Kota Bukittinggi

Hasil ini dapat membantu dan menjadi bahan pendukung pihak Puskesmas Guguk Panjang dalam mengambil kebijakan.