

**KLASIFIKASI FAKTOR-FAKTOR PENYEBAB PENYAKIT
DIABETES MELLITUS DI RSUD KOTA PARIAMAN TAHUN 2022
MENGUNAKAN ALGORITMA C4.5**

TUGAS AKHIR

Diajukan Sebagai Salah Satu Persyaratan Untuk Memperoleh Gelar Ahli Madya



Oleh

LATHIFAH AZ-ZAHRA

NIM 20037033

**PROGRAM STUDI DIPLOMA III STATISTIKA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS NEGERI PADANG
2023**

PENGESAHAN LULUS UJIAN TUGAS AKHIR




Nama : Lathifah Az-Zahra
NIM : 20037033
Program Studi : D3 Statistika
Departemen : Statistika
Fakultas : Matematika dan Ilmu Pengetahuan Alam

KLASIFIKASI FAKTOR-FAKTOR PENYEBAB PENYAKIT DIABETES MELLITUS DI RSUD KOTA PARIAMAN TAHUN 2022 MENGUNAKAN ALGORITMA C4.5

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

Padang, 07 November 2023

Tim Penguji

	Nama	Tanda Tangan
Ketua	: Dr. Syafriandi, M.Si	
Anggota	: Dr. Donny Permana, M.Si	
Anggota	: Zilrahmi, M.Si	

PERSETUJUAN TUGAS AKHIR

**KLASIFIKASI FAKTOR-FAKTOR PENYEBAB PENYAKIT
DIABETES MELLITUS DI RSUD KOTA PARIAMAN TAHUN 2022
MENGUNAKAN ALGORITMA C4.5**

Nama : Lathifah Az-Zahra
NIM : 20037033
Program Studi : D3 Statistika
Departemen : Statistika
Fakultas : Matematika dan Ilmu Pengetahuan Alam

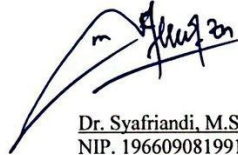
Padang, 07 November 2023

Mengetahui:
Kepala Departemen Statistika

Disetujui Oleh:
Pembimbing



Dr. Yenni Kurniawati, S.Si., M.Si
NIP. 198402232010122005



Dr. Syafriandi, M.Si
NIP. 196609081991031003

SURAT PERNYATAAN TIDAK PLAGIAT

Saya yang bertandatangan di bawah ini:

Nama : Lathifah Az-Zahra
NIM : 20037033
Program Studi : D3 Statistika
Departemen : Statistika
Fakultas : Matematika dan Ilmu Pengetahuan Alam

Dengan ini menyatakan bahwa, skripsi saya dengan judul **“Klasifikasi Faktor-Faktor Penyebab Penyakit Diabetes Mellitus di RSUD Kota Pariaman Tahun 2022 Menggunakan 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.

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. 19840223201012205

Saya yang menyatakan,



Lathifah Az-Zahra
NIM. 20037033

**KLASIFIKASI FAKTOR-FAKTOR PENYEBAB PENYAKIT DIABETES
MELLITUS DI RSUD KOTA PARIAMAN TAHUN 2022
MENGUNAKAN ALGORITMA C4.5**

ABSTRAK

Lathifah Az - Zahra

Diabetes mellitus adalah kondisi ketika kadar glukosa darah mengalami peningkatan. Diabetes mellitus merupakan salah satu penyakit tidak menular. Berdasarkan Riset kesehatan dasar Provinsi Sumatera Barat 2018, prevalensi penderita diabetes mellitus di Sumatera Barat sebesar 1,6%. Sedangkan, prevalensi penderita diabetes mellitus di Kota Pariaman sebesar 3,4%, yang mana angka ini melebihi angka prevalensi diabetes mellitus Provinsi Sumatera Barat. Berdasarkan Kemenkes RI (2018) diabetes mellitus dapat memicu penyakit-penyakit lainnya seperti serangan jantung, stroke, kebutaan dan gagal ginjal bahkan menyebabkan kelumpuhan dan kematian. Hal yang dapat dilakukan adalah menentukan ciri-ciri penderita diabetes mellitus dan mengetahui tingkat akurasi hasil klasifikasi pasien yang menderita penyakit diabetes mellitus berdasarkan faktor-faktor penyebabnya.

Penelitian ini merupakan penelitian terapan, yaitu penelitian yang dilakukan berdasarkan penerapan dalam menyelesaikan suatu permasalahan. Data yang digunakan pada penelitian ini adalah data pasien poliklinik penyakit dalam, berdasarkan penyebab diabetes mellitus dari pasien diabetes mellitus di RSUD Kota Pariaman. Variabel yang digunakan pada penelitian ini adalah Jenis Kelamin (X1), Umur (X2), Glukosa (X3), Kolesterol (X4), HDL (X5), LDL (X6), Trigliserida (X7), dan Kriteria pasien (Y). Langkah yang dilakukan yaitu memisah data latih dan data uji sebanyak 80% dan 20%. Dalam Algoritma C4.5 pembentukan model pohon keputusan menggunakan nilai *entropy* dan *information gain*. Model pohon keputusan yang telah diperoleh dilakukan validasi sistem untuk mengukur kinerjanya dengan menggunakan *confusion matrix*.

Hasil penelitian didapat bahwa variabel Glukosa yang menjadi penciri utama pasien menderita diabetes mellitus. Dari semua variabel yang digunakan terbukti satu diantaranya tidak mempengaruhi variabel terikat yaitu variabel HDL. Didapat nilai akurasi sebesar 97,9%, artinya kualitas klasifikasi yang dilakukan berada pada tingkat “sangat tinggi”.

Kata Kunci: Diabetes Mellitus, Faktor Penyebab, Pohon Keputusan

CLASSIFICATION OF FACTORS INCLUDING MELLITUS DIABETES IN THE 2022 CITY RATES USING C4.5 ALGORITM

ABSTRACT

Lathifah Az-Zahra

Diabetes mellitus is a condition when blood glucose levels are elevated. Diabetes mellitus is one of the non-communicable diseases. According to the Basic Health Survey of the West Sumatra Province in 2018, the prevalence of diabetes mellitus in the Western Sumatra province was 1.6 per cent, whereas in the City of Pariaman it was 3.4 per cent. According to Kemenkes RI (2018) diabetes mellitus can trigger other diseases such as heart attacks, strokes, blindness and kidney failure even lead to paralysis and death. What can be done is to determine the characteristics of people with diabetes mellitus and to know the degree of accuracy of the classification of patients suffering from diabetes Mellitus based on the factors that cause it.

This research is applied research, that is, research that is carried out on the basis of application in solving a problem. The data used in this study are data from patients in poliklinics of internal diseases, based on the causes of diabetes mellitus from diabetic patients in RSUD Kota Pariaman. The variables used in this study are Gender type (X1), Age (X2), Glucose (X3), Cholesterol (X4), HDL (X5), LDL(X6), Triglycerides (X7), and Patient criteria (Y). The step is to separate the training data from the test data by 80% and 20%. In algorithm C4.5, the decision tree model is formed using entropy and information gain values.

The results of the study found that the glucose variable is the primary detector of patients suffering from diabetes mellitus. Of all the variables used, one of them proved not to affect the bound variable HDL. The accuracy value was 97.9%, which means that the quality of the classification was at a “very high” level.

Keywords: Diabetes Mellitus, Causal Factors, Decision Tree

KATA PENGANTAR

Alhamdulillah puji syukur penulis panjatkan kepada Allah SWT Tuhan semesta alam yang telah melimpahkan rahmat, hidayah-Nya kepada penulis, sehingga penulis dapat menyelesaikan Tugas Akhir dengan judul “**Klasifikasi Faktor-Faktor Penyebab Penyakit Diabetes Mellitus di RSUD Kota Pariaman Tahun 2022 Menggunakan 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 FMIPA Universitas Negeri Padang. 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:

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., dan Ibu Zilrahmi, M.Si., dosen pembahas yang telah memberikan arahan bimbingan, dan nasehat kepada penulis dalam menyelesaikan Tugas Akhir.
3. Ibu Dr. Yenni Kurniawati, M.Si., Kepala Departemen Statistika yang telah memberikan arahan bimbingan dan nasehat kepada penulis dalam menyelesaikan Tugas Akhir.
4. Bapak Dodi Vionanda, M.Si, Ph.D., Kepala Prodi Diploma III Departemen Statistika FMIPA Universitas Negeri Padang

5. Bapak dan Ibu Dosen, Staf Pengajar dan Karyawan Departemen Statistika FMIPA Universitas Negeri Padang yang telah membimbing dan berbagi ilmu kepada penulis selama perkuliahan.
6. Pemimpin dan Staf RSUD Kota Pariaman yang telah mengizinkan penulis melakukan penelitian.
7. Kedua orang tua penulis yang telah mendukung dan membantu penulis dalam menyelesaikan Tugas Akhir baik moril dan dukungan lainnya.

Semoga bantuan dan bimbingan yang diberikan kepada penulis menjadi amal ibadah dan dibalas 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 2023

Lathifah Az - Zahra

DAFTAR ISI

ABSTRAK	i
KATA PENGANTAR	iii
DAFTAR ISI	iv
DAFTAR TABEL	vii
DAFTAR GAMBAR	viii
DAFTAR LAMPIRAN	ix
BAB I PENDAHULUAN	1
A. Latar Belakang Masalah.....	1
B. Batasan Masalah.....	7
C. Rumusan Masalah.....	7
D. Tujuan Penelitian.....	8
E. Manfaat Penelitian.....	8
BAB II LANDASAN TEORI	10
A. Diabetes Mellitus.....	10
B. Klasifikasi Diabetes Mellitus.....	11
C. Faktor-Faktor Penyebab Diabetes Mellitus.....	12
D. Data Mining.....	15
E. Data Training dan Data Testing.....	15
F. Klasifikasi.....	16
G. Decision Tree.....	17
H. Algoritma C4.5.....	18
I. Validasi Model.....	21
BAB III METODOLOGI PENELITIAN	25
A. Jenis Penelitian.....	25
B. Jenis dan Sumber Data.....	25
C. Populasi dan Sampel.....	25
D. Prosedur Penelitian.....	26
E. Teknik Analisis Data.....	27
BAB IV HASIL DAN PEMBAHASAN	29
A. Hasil Penelitian.....	29
B. Pembahasan.....	48
BAB V PENUTUP	50

A. Kesimpulan	50
B. Saran	52
DAFTAR PUSTAKA.....	53
LAMPIRAN	56

DAFTAR TABEL

Tabel	Halaman
1. <i>Confusion Matrix</i>	21
2. Kriteria Kualitas Akurasi	22
3. Kriteria Kualitas <i>Precision</i>	23
4. Kriteria Kualitas <i>Recall</i>	24
5. Variabel Penelitian	26
6. Struktur Data	27
7. Sampel Data Pembuatan Model Pohon Keputusan Algoritma C4.5.....	35
8. Hasil Perhitungan <i>Entropy</i> dan <i>Information Gain</i>	44
9. <i>Confusion Matrix</i> Klasifikasi Kriteria Diabetes Mellitus	46

DAFTAR GAMBAR

Gambar	Halaman
Gambar 1. Prevelensi Penderita Diabetes Mellitus di Indonesia dan Sumatera Barat Tahun 2007, 2013, dan Tahun 2018.....	2
Gambar 2. Prevelensi Penderita Diabetes Mellitus di Kota/Kab.Provinsi Sumatera Barat Tahun 2018.	3
Gambar 3. Prevelensi 10 Penyakit Dalam di RSUD Kota Pariaman Tahun 2022.	4
Gambar 4. Proses Klasifikasi	16
Gambar 5. Model Pohon Keputusan	17
Gambar 6. Flowchart algoritma C4.5	28
Gambar 7. Jumlah Pasien berdasarkan Jenis Kelamin dan Kriteria Pasien	29
Gambar 8. Jumlah Pasien berdasarkan Umur dan Kriteria Pasien.....	30
Gambar 9. Jumlah Pasien berdasarkan Glukosa dan Kriteria Pasien.....	31
Gambar 10. Jumlah Pasien berdasarkan Kolesterol dan Kriteria Pasien.....	32
Gambar 11. Jumlah Pasien berdasarkan HDL dan Kriteria Pasien.....	32
Gambar 12. Jumlah Pasien berdasarkan LDL dan Kriteria Pasien	33
Gambar 13. Jumlah Pasien berdasarkan Trigliserida dan Kriteria Pasien.....	34
Gambar 14. Tahapan pembentukan model pohon keputusan algoritma C4.5.....	34
Gambar 15. Model Pohon Keputusan Algoritma C4.5.....	45

DAFTAR LAMPIRAN

Lampiran	Halaman
1. Sampel Data Kriteria Pasien di RSUD Kota Pariaman.....	56
2. Output Data <i>Training</i>	57
3. Output Data <i>Testing</i>	58
4. Output Model Pohon Keputusan Menggunakan <i>Rapidminer</i>	59
5. Output <i>Confusion Matrix</i> Model Pohon Keputusan Menggunakan <i>Rapidminer</i>	60

BAB I

PENDAHULUAN

A. Latar Belakang Masalah

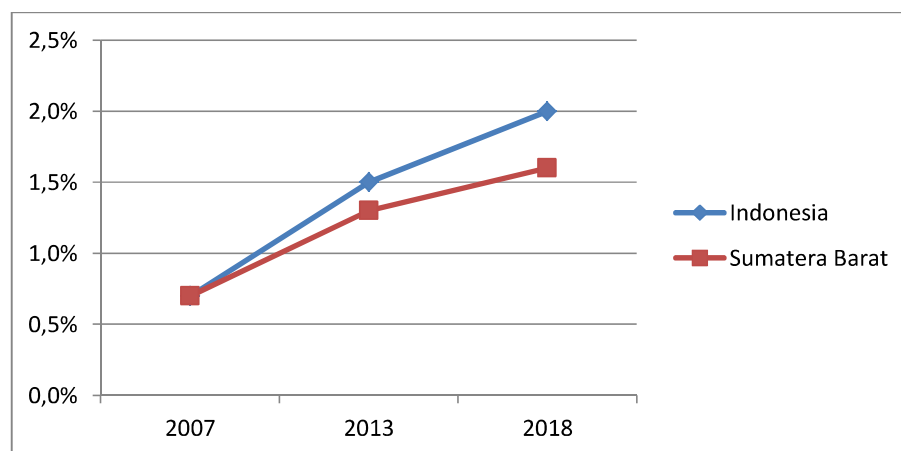
Diabetes Mellitus (DM) merupakan salah satu penyakit tidak menular yang menjadi permasalahan di Indonesia. Kemenkes RI (2020) menyatakan penyebab DM secara umum terjadi karena pola hidup yang tidak sehat, seperti kurangnya mengonsumsi sayur dan buah serta lebih banyak mengonsumsi makanan cepat saji (*junk food*) dan makanan tinggi gula. Menurut American Diabetes Association (2018) DM adalah penyakit kronis yang kompleks, secara medis terjadi karena kelainan sekresi insulin, kerja insulin atau keduanya, sehingga menyebabkan tubuh mengalami gangguan dalam metabolisme karbohidrat, protein, dan lemak yang berdampak pada kadar glukosa darah yang cenderung meningkat atau hiperglikemia.

Kemenkes RI (2020) menyatakan bahwa DM digolongkan menjadi tiga jenis, yaitu DM tipe 1, tipe 2, dan diabetes gestasional. Menurut International Diabetes Federation (2019) DM tipe 1 dicirikan dengan reaksi *autoimun* yang menyebabkan sistem kekebalan tubuh menyerang sel beta pada pankreas sehingga tidak bisa memproduksi insulin sama sekali. Sedangkan DM tipe 2 terjadi akibat adanya resistensi insulin, yang mana sel-sel dalam tubuh tidak mampu merespon sepenuhnya insulin, dan diabetes gestasional disebabkan karena naiknya berbagai kadar hormon saat hamil yang bisa menghambat kerja insulin.

Penyakit DM dapat menyebabkan gangguan metabolik menahun yang lebih dikenal sebagai pembunuh manusia secara diam-diam atau "*silent killer*". Seringkali penyandanginya tidak menyadari telah menderita penyakit DM, dan

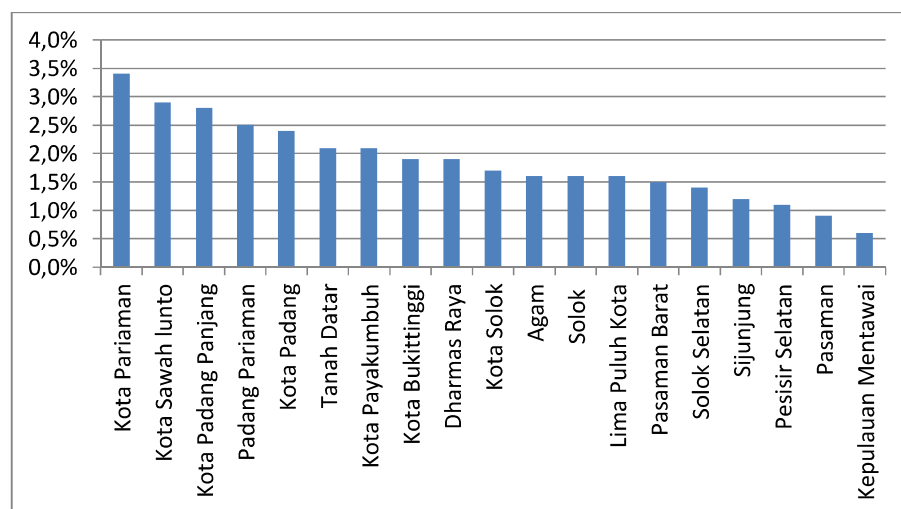
mengalami keterlambatan dalam menanganinya, sehingga banyak terjadi komplikasi. Menurut Kemenkes RI (2018) DM juga dikenal sebagai “*Mother of Disease*” karena dapat memicu penyakit-penyakit lainnya seperti serangan jantung, stroke, kebutaan dan gagal ginjal bahkan menyebabkan kelumpuhan dan kematian. Apabila dibiarkan tidak terkendali maka penyakit ini dapat menimbulkan komplikasi lain yang membahayakan kesehatan.

Berdasarkan hasil Riset Kesehatan Dasar Nasional (Riskesdas) yang dilakukan setiap 5 tahun sekali, terlihat bahwa angka prevelensi penderita DM di Indonesia mencapai 0,6% pada tahun 2007. Angka ini menunjukkan peningkatan dari hasil Riskesdas setelahnya, yang mana angka prevelensi penderita DM pada tahun 2013 sebesar 1,5% dan pada tahun 2018 sebesar 2,0%. Sementara itu, prevelensi penderita DM di Sumatera Barat mencapai 0,6% pada tahun 2007. Angka ini menunjukkan peningkatan dari hasil Riskesdas setelahnya, yang mana angka prevelensi penderita DM pada tahun 2013 sebesar 1,3% dan pada tahun 2018 sebesar 1,6%. Gambar 1 menampilkan grafik prevelensi penderita DM di Indonesia dan Provinsi Sumatera Barat Tahun 2007, 2013, dan 2018.



Gambar 1. Prevelensi Penderita Diabetes Mellitus di Indonesia dan Sumatera Barat Tahun 2007, 2013, dan Tahun 2018

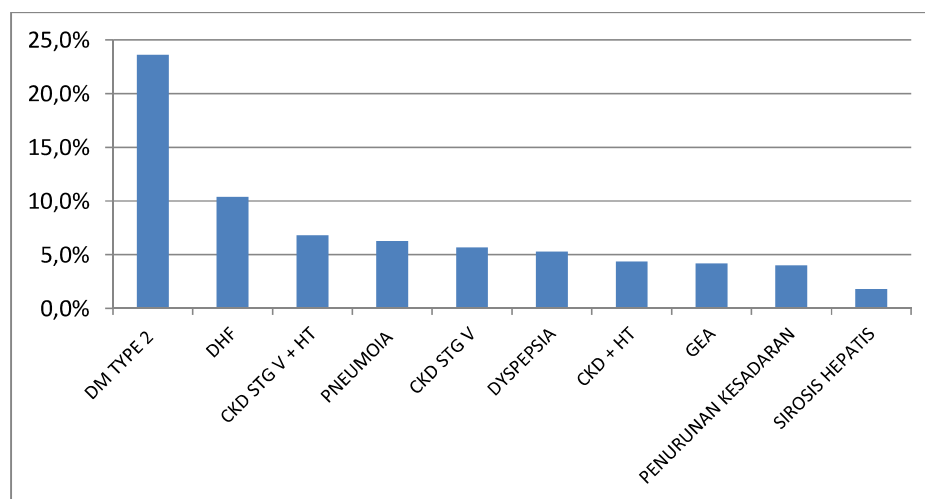
Hasil Riskesdas Provinsi Sumatera Barat tahun 2018, menunjukkan angka prevalensi penderita diabetes di Sumatera Barat sebesar 1,6%. Angka prevalensi penderita diabetes posisi pertama ditempati oleh Kota Pariaman sebesar 3,4% dari 19 Kota/Kabupaten yang ada di Provinsi Sumatera Barat, yang mana angka ini melebihi angka prevalensi penderita DM Provinsi Sumatera Barat. Gambar 2 menampilkan grafik prevalensi penderita DM di Kota/Kabupaten Provinsi Sumatera Barat.



Gambar 2. Prevelensi Penderita Diabetes Mellitus di Kota/Kab.Provinsi Sumatera Barat Tahun 2018.

Berdasarkan Gambar 2, prevalensi penderita DM di Sumatera Barat tahun 2018 secara signifikan dapat tergantung pada kondisi ekonomi dan tingkat kesehatan suatu wilayah. Dalam kawasan perkotaan dengan tingkat ekonomi yang tinggi menunjukkan prevalensi penderita DM yang tinggi, artinya kota di Provinsi Sumatera Barat memiliki tingkat kesehatan yang rendah. Sedangkan, di kawasan kabupaten dengan tingkat ekonomi yang rendah menunjukkan prevalensi penderita DM yang rendah, artinya kabupaten di Provinsi Sumatera Barat memiliki tingkat kesehatan yang tinggi.

Berdasarkan data rekam medis pasien Poliklinik Penyakit Dalam di RSUD Kota Pariaman tahun 2022, menunjukkan angka prevelensi penyakit tertinggi adalah DM Tipe II sebesar 23,6% yang mendapatkan penanganan medis di rumah sakit tersebut. Artinya, penyakit DM Tipe II ini mendominasi kasus yang terjadi pada pasien yang berobat di Poliklinik Penyakit Dalam RSUD Kota Pariaman. Gambar 3 menampilkan grafik prevelensi 10 penyakit dalam di RSUD Kota Pariaman.



Gambar 3. Prevelensi 10 Penyakit Dalam di RSUD Kota Pariaman Tahun 2022.

Menurut Parkeni (2015) tingginya kejadian DM tipe 2 pada masyarakat disebabkan oleh glukosa darah yang memiliki peranan terhadap kenaikan kadar lemak darah, peningkatan kadar trigliserida terutama karena hubungannya dengan kadar kolesterol *low density lipoprotein* (LDL) tinggi atau resistensi insulin, kolesterol total serta penurunan kolesterol *high density lipoprotein* (HDL). Peningkatan glukosa, trigliserida, LDL, kolesterol dan penurunan level HDL merupakan karakteristik timbulnya penyakit DM.

Upaya pencegahan dan pengendalian penyakit tidak menular yang sedang dikembangkan di Indonesia adalah Pos pembinaan terpadu penyakit tidak

menular (Posbindu PTM). Namun, menurut pusat data dan informasi Kementerian Kesehatan RI tingkat ketidakpatuhan penderita DM dalam kontrol gula darah masih memiliki angka yang cukup tinggi pada tahun 2018. Kemenkes RI (2020) membuktikan dengan data laporan konsumsi makanan dan minuman manis, yang mana 47,8% survei yang dilakukan terhadap penyandang DM mengonsumsi makanan dan minuman manis 1–6 kali/minggu dan hanya 12% mengonsuminya kurang dari 3 kali perbulan.

Oleh karena itu, diperlukan analisis terhadap data penyakit yang ada berdasarkan faktor-faktor penyebabnya. Salah satu analisis yang dapat digunakan adalah data mining untuk menemukan pengetahuan dari data yang berjumlah besar. Menurut Sartika dkk., (2019) Data mining adalah suatu proses untuk memperoleh informasi penting yang sangat berguna dari informasi yang sebelumnya tidak diketahui dan pengetahuan yang diperoleh dapat dimengerti dari sebuah data. Untuk bisa mendapatkan informasi dari sekumpulan *database* yang besar diperlukan proses data mining seperti klasifikasi. Menurut Andriani (2018) Klasifikasi adalah proses menempatkan suatu objek kedalam satu set kategori berdasarkan objek yang bersangkutan.

Metode klasifikasi pada penelitian ini digunakan untuk membantu dalam memahami pengelompokan data pasien yang menderita penyakit DM, berdasarkan faktor-faktor penyebab penyakit DM. Menurut Arisman (2014) faktor-faktor penyebab penyakit DM adalah jenis kelamin, umur, glukosa, kolesterol, HDL, LDL, dan trigliserida. Hasil klasifikasi ini akan mengidentifikasi pola-pola yang signifikan dalam data, sehingga memungkinkan pengelompokan pasien ke dalam kategori risiko yang berbeda.

Adapun metode dalam klasifikasi antara lain *Decision Tree*, *Naive Bayes*, *Support Vector Machine*, dan *Random Forest*. Salah satu metode yang dapat digunakan yaitu *Decision Tree* atau pohon keputusan. *Decision tree* adalah suatu metode klasifikasi yang banyak digunakan karena kemudahannya untuk diinterpretasikan. Pohon keputusan memiliki model seperti namanya yaitu struktur pohon atau struktur berhirarki. Konsep dari pohon keputusan adalah mengubah sejumlah data menjadi suatu aturan keputusan. Pohon keputusan juga berguna dalam pengeksplorasian data, menemukan suatu hubungan antara sejumlah variabel dengan variabel target. Menurut Bisri (2015) *decision tree* memiliki kelebihan seperti dapat memecahkan masalah *overfitting*, menangani nilai atribut yang hilang pada data latih dan dapat meningkatkan efisiensi komputasi.

Algoritma C4.5 termasuk dalam algoritma klasifikasi data dengan teknik pohon keputusan yang dapat menangani atribut bertipe diskret dan numerik, menghasilkan aturan-aturan yang mudah diinterpretasikan dan tercepat diantara algoritma-algoritma yang lain yaitu *ID3*, *CART*, *CHAID*, dan *Random Forest*. Dalam melakukan proses klasifikasi, algoritma ini membentuk pohon keputusan yang merupakan sebuah struktur yang berfungsi menggolongkan data sesuai aturan pengelompokan sampai pada bagian terkecil. Menurut Wilem dkk., (2021) Algoritma C4.5 digunakan untuk mengenali pola dari dataset menggunakan sebuah atribut sebagai pengenal terhadap karakteristik data, dimana atribut tersebut pada umumnya diberi nama label atau kelas.

Beberapa penelitian terkait yang berhubungan dengan penelitian ini dilakukan oleh Yusa dkk., (2015) menggunakan model *decision tree* dengan algoritma C4.5

untuk klasifikasi terjadinya obesitas, Rohman (2013) dan Ente dkk., (2019) juga telah menggunakan metode algoritma C4.5 untuk memprediksi penyakit jantung dan penyakit DM. Oleh karena itu dalam penelitian ini akan digunakan algoritma C4.5 untuk memprediksi status penyakit dan mengklasifikasi faktor-faktor yang menjadi penyebab penyakit DM.

Berdasarkan latar belakang di atas, penulis melakukan penelitian yang berjudul **“Klasifikasi Faktor-Faktor Penyebab Penyakit Diabetes Mellitus di RSUD Kota Pariaman Tahun 2022 Menggunakan Algoritma C4.5”**.

B. Batasan Masalah

Batasan masalah penelitian ini adalah sebagai berikut.

1. Metode klasifikasi yang digunakan adalah metode *decission tree* dengan algoritma C4.5
2. Data yang digunakan adalah data pasien poli penyakit dalam di RSUD Kota Pariaman Tahun 2022.

C. Rumusan Masalah

Rumusan masalah pada penelitian ini adalah sebagai berikut.

1. Bagaimana ciri-ciri klasifikasi pasien yang menderita penyakit diabetes mellitus berdasarkan faktor-faktor penyebab penyakit diabetes mellitus di RSUD Kota Pariaman menggunakan algoritma C4.5?
2. Bagaimana tingkat akurasi hasil klasifikasi pasien yang menderita penyakit diabetes mellitus berdasarkan faktor-faktor penyebab penyakit diabetes mellitus menggunakan algoritma C4.5 yang diperoleh?

D. Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut.

1. Untuk menentukan ciri-ciri klasifikasi pasien yang menderita penyakit diabetes mellitus berdasarkan faktor-faktor penyebab penyakit diabetes mellitus di RSUD Kota Pariaman dengan menggunakan algoritma C4.5.
2. Untuk mengetahui tingkat akurasi hasil klasifikasi pasien yang menderita penyakit diabetes mellitus berdasarkan faktor-faktor penyebab penyakit diabetes mellitus menggunakan algoritma C4.5 yang diperoleh.

E. Manfaat Penelitian

1. Bagi Penulis

Untuk memenuhi persyaratan memperoleh gelar ahli madya jurusan Statistika Universitas Negeri Padang. Tugas akhir ini sebagai penerapan ilmu yang telah dipelajari selama perkuliahan.

2. Bagi Pembaca

Penelitian ini bisa dijadikan sebagai referensi dan bahan acuan bagi peneliti selanjutnya dalam menerapkan metode *decission tree* dengan algoritma C4.5.

3. Bagi RSUD Kota Pariaman

Hasil dari penelitian ini berupa model dan aturan/*rule* untuk mengklasifikasikan berdasarkan faktor-faktor penyebab penyakit diabetes mellitus yang nantinya bisa membantu pihak RSUD Kota Pariaman dalam pengambilan keputusan pasien diabetes mellitus.

4. Bagi Pasien Diabetes Mellitus

Penelitian mengenai faktor-faktor penyebab diabetes mellitus dapat membantu pasien mengambil keputusan yang lebih tepat dalam pengelolaan kesehatan sehari-hari.

5. Bagi Masyarakat

Penelitian ini dapat meningkatkan kesadaran masyarakat tentang faktor-faktor risiko diabetes mellitus. Dengan mengetahui penyebab-penyebab potensial penyakit ini, masyarakat dapat meningkatkan pola hidup yang lebih sehat sehingga masyarakat dapat membantu mengurangi jumlah kasus diabetes mellitus.