

## BAB II LANDASAN TEORI

### 2.1 *Data mining*

*Data mining* adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam *database*. *Data mining* adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait besar. *Data mining* adalah serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual. *Data mining* merupakan bidang dari beberapa bidang keilmuan yang menyatukan teknik dari pembelajaran mesin, pengenalan pola, statistik, *database*, dan visualisasi untuk penanganan permasalahan pengambilan informasi dari *database* yang besar. Dari definisi-definisi yang telah disampaikan hal yang terpenting yang terkait dengan *Data mining* adalah:

1. *Data mining* merupakan suatu proses otomatis terhadap data yang sudah ada.
2. Data yang diproses berupa data yang sangat besar.
3. Tujuan *Data mining* adalah mendapatkan hubungan atau pola yang mungkin memberikan indikasi yang bermanfaat.

Beberapa definisi awal dari *Data mining* menyertakan fokus pada proses otomatisasi. *Data mining* otomatisasi tidak menggantikan campur tangan manusia. Manusia harus ikut aktif dalam setiap fase dalam proses *Data mining*. Kehebatan kemampuan algoritma *Data mining* yang terdapat dalam perangkat lunak analisis yang terdapat saat ini memungkinkan terjadinya kesalahan pengguna yang berakibat fatal. Pengguna mungkin menerapkan analisis yang tidak tepat terhadap kumpulan data dengan menggunakan pendekatan yang berbeda. Oleh karenanya, dibutuhkan

pemahaman tentang statistik dan struktur model matematika yang mendasari kerja perangkat lunak (Kusrini; Emha Taufiq Lutfi, 2009).

## 2.2 Pengelompokan *Data mining*

*Data mining* dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan, yaitu :

### 1. Deskripsi

Terkadang peneliti dan analisis secara sederhana ingin mencoba mencari cara untuk menggambarkan pola dan kecenderungan yang terdapat dalam data. Sebagai contoh, petugas pengumpulan suara mungkin tidak dapat menemukan keterangan atau fakta bahwa siapa yang tidak cukup profesional akan sedikit didukung dalam pemilihan presiden. Deskripsi dari pola dan kecenderungan sering memberikan kemungkinan penjelasan untuk suatu pola atau kecenderungan.

### 2. Estimasi

Estimasi hampir sama dengan klasifikasi, kecuali variabel target estimasi lebih ke arah numerik daripada ke arah kategori. Model dibangun menggunakan *record* lengkap yang menyediakan nilai dari variabel target sebagai nilai prediksi. Selanjutnya, pada peninjauan berikut estimasi nilai variabel target dibuat berdasarkan nilai variabel prediksi. Sebagai contoh, akan dilakukan estimasi tekanan darah sistolik pada pasien rumah sakit berdasarkan umur pasien, jenis kelamin, indeks berat badan, dan level sodium darah. Hubungan antar tekanan darah sistolik dan nilai variabel prediksi dalam proses pembelajaran akan menghasilkan model estimasi. Model estimasi yang dihasilkan dapat digunakan untuk kasus baru lainnya.

### 3. Prediksi

Prediksi hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan ada di masa mendatang.

Contoh prediksi dalam bisnis dan penelitian adalah :

- Prediksi harga beras dalam tiga bulan yang akan datang.

- Prediksi presentase kenaikan lalu lintas tahun depan jika batas bawah kecepatan dinaikkan.

Beberapa metode dan teknik yang digunakan dalam klasifikasi dan estimasi dapat pula digunakan (untuk keadaan yang tepat) untuk prediksi.

#### 4. Klasifikasi

Dalam klasifikasi, terdapat target variabel kategori. Sebagai contoh, penggolongan pendapatan dapat dipisahkan dalam tiga kategori, yaitu pendapatan tinggi, pendapatan sedang, dan pendapatan rendah.

Contoh lain klasifikasi dalam bisnis dan penelitian adalah :

- Menentukan apakah suatu transaksi kartu kredit merupakan transaksi yang curang atau bukan.
- Memperkirakan apakah suatu pengajuan hipotek oleh nasabah merupakan suatu kredit yang baik atau buruk.
- Mendiagnosis penyakit seorang pasien untuk mendapatkan termasuk kategori penyakit apa.

#### 5. Pengklusteran

Pengklusteran merupakan pengelompokan *record*, pengamatan, atau memperhatikan dan membentuk kelas objek-objek yang memiliki kemiripan. Kluster adalah kumpulan *record* yang memiliki kemiripan satu dengan yang lainnya dan memiliki ketidakmiripan dengan *record-record* dalam kluster lain.

Pengklusteran berbeda dengan klasifikasi yaitu tidak adanya variabel target dalam pengklusteran. Pengklusteran tidak mencoba untuk melakukan klasifikasi, mengestimasi, atau memprediksi nilai dari variabel target. Akan tetapi, algoritma pengklusteran mencoba untuk melakukan pembagian terhadap keseluruhan data menjadi kelompok-kelompok yang memiliki kemiripan (homogen), yang mana kemiripan *record* dalam satu kelompok lain akan bernilai minimal.

Contoh pengklusteran dalam bisnis dan penelitian adalah :

- Mendapatkan kelompok-kelompok konsumen untuk target pemasaran dari suatu produk bagi perusahaan yang tidak memiliki dana pemasaran yang besar.
- Untuk tujuan audit akutansi, yaitu melakukan pemisahan terhadap perilaku finansial dalam baik dan mencurigakan.
- Melakukan pengklusteran terhadap ekspresi dari gen, untuk mendapatkan kemiripan perilaku dari gen dalam jumlah besar.

#### 6. Asosiasi

Tugas asosiasi dalam *Data mining* adalah menemukan atribut yang muncul dalam satu waktu. Dalam dunia bisnis lebih umum disebut analisis keranjang belanja.

Contoh asosiasi dalam bisnis dan penelitian adalah :

- Meneliti jumlah pelanggan dari perusahaan telekomunikasi seluler yang diharapkan untuk memberikan respon positif terhadap penawaran *upgrade* layanan yang diberikan.
- Menemukan barang dalam supermarket yang dibeli secara bersamaan dan barang yang tidak pernah dibeli secara bersamaan.

### 2.3 Proses *Data Mining*

Secara sistematis, ada tiga langkah utama dalam data mining :

#### 1) Eksplorasi atau pemrosesan awal data

Eksplorasi atau pemrosesan awal data terdiri dari pembersihan data, normalisasi data, transformasi data, penanganan data yang salah, reduksi dimensi, pemilihan subset fitur, dan sebagainya.

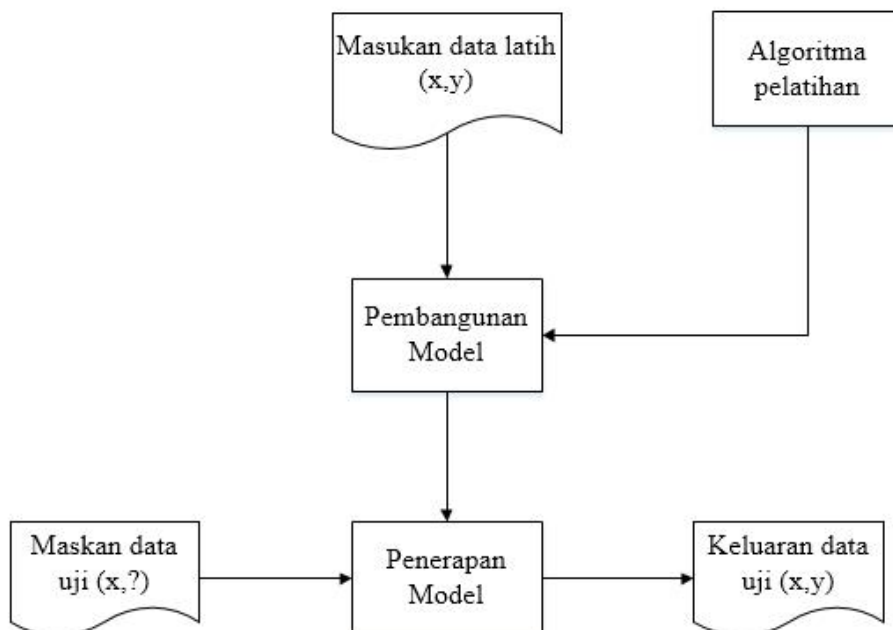
#### 2) Membangun model dan melakukan validasi terhadapnya

Berarti melakukan analisis berbagai model dan memilih model dengan kinerja prediksi yang terbaik. Dalam langkah ini digunakan metode-metode seperti klasifikasi, regresi, analisis cluster, deteksi anomali, analisis asosiasi, analisis pola sekuensial, dan sebagainya. Dalam beberapa referensi, deteksi anomali juga masuk dalam langkah eksplorasi. Akan

tetapi deteksi anomali juga dapat digunakan sebagai algoritma utama, terutama untuk mencari data-data yang spesial (Prasetyo, 2014).

## 2.4 Konsep Klasifikasi

Klasifikasi dapat didefinisikan secara detail sebagai suatu pekerjaan yang melakukan pelatihan / pembelajaran terhadap fungsi target  $f$  yang memetakan setiap vektor (set fitur)  $x$  ke dalam suatu dari sejumlah label kelas  $y$  yang tersedia (Prasetyo, 2014). Klasifikasi merupakan suatu model yang menerima masukan kemudian mampu melakukan pemikiran terhadap masukan tersebut dan memberikan jawaban sebagai keluaran hasil pemikirannya. Kerangka kerja klasifikasi ditunjukkan pada Gambar 2.1. Pada gambar tersebut, disediakan sejumlah data latih  $(x,y)$  untuk digunakan sebagai data pembangunan model, kemudian menggunakan model tersebut untuk memprediksi kelas dari data uji  $(x,?)$  sehingga data uji  $(x,?)$  diketahui kelas  $y$  yang seharusnya.



Gambar 2.1 Proses pekerjaan klasifikasi

Model yang sudah dibangun pada saat pelatihan kemudian dapat digunakan untuk memprediksi label kelas dari data baru yang belum diketahui label kelasnya. Dalam pembangunan model selama proses pelatihan tersebut diperlukan adanya suatu algoritma untuk membangunnya yang disebut sebagai algoritma pelatihan (*learning algorithm*). Kerangka kerja seperti yang ditunjukkan pada Gambar 2.1 meliputi dua langkah proses yaitu induksi dan deduksi. Induksi merupakan suatu langkah untuk membangun model klasifikasi dari data latih yang diberikan, disebut juga proses pelatihan. Sedangkan deduksi merupakan suatu langkah untuk menerapkan model tersebut pada data uji sehingga data uji dapat diketahui kelas yang sesungguhnya atau disebut juga proses prediksi.

## 2.5 Pohon Keputusan (*Decision tree*)

### 2.5.1 Pengertian Pohon Keputusan

Pohon keputusan merupakan metode klasifikasi dan prediksi yang sangat kuat dan terkenal. Metode pohon keputusan mengubah fakta yang sangat besar menjadi pohon keputusan yang mempresentasikan aturan. Aturan dapat dengan mudah dipahami dengan bahasa alami. Dan mereka juga dapat diekspresikan dalam bentuk bahasa basis data seperti *Structured Query Language* untuk mencari *record* pada kategori tertentu. Pohon keputusan juga berguna untuk mengeksplorasi data, menemukan hubungan tersembunyi antara sejumlah calon variabel input dengan sebuah variabel target. Karena pohon keputusan memadukan antara eksplorasi data dan pemodelan, sangat bagus sebagai langkah awal dalam proses pemodelan bahkan ketika dijadikan sebagai model akhir dari beberapa teknik lain.

Sebuah pohon keputusan adalah sebuah struktur yang dapat digunakan untuk membagi kumpulan data yang besar menjadi himpunan-himpunan *record* yang lebih kecil dengan menerapkan serangkaian

aturan keputusan. Dengan masing-masing rangkaian pembagian, anggota himpunan hasil menjadi mirip satu dengan yang lain.

Sebuah model pohon keputusan terdiri dari sekumpulan aturan untuk membagi sejumlah populasi yang heterogen menjadi lebih kecil, lebih *homogen* dengan memperhatikan pada variabel tujuannya. Sebuah pohon keputusan mungkin dibangun dengan seksama secara manual atau dapat tumbuh secara otomatis dengan menerapkan salah satu atau beberapa algoritma pohon keputusan untuk memodelkan himpunan data yang belum terklarifikasi.

Variabel tujuannya biasanya dikelompokkan dengan pasti dan model pohon keputusan lebih mengarah pada perhitungan probabilitas dari tiap-tiap *record* terhadap kategori-kategori tersebut atau untuk mengklasifikasi *record* dengan mengelompokkannya dalam satu kelas. Pohon keputusan juga dapat digunakan untuk mengetimasi nilai dari variabel *continue* meskipun ada beberapa teknik yang lebih sesuai untuk kasus ini.

Banyak algoritma yang dapat dipakai dalam pembentukan pohon keputusan, antara lain ID3, CART, dan C4.5. Dalam data pohon keputusan biasanya dinyatakan dalam bentuk tabel dengan atribut dan *record*. Atribut menyatakan suatu parameter yang dibuat sebagai kriteria dalam pembentukan pohon. Misalkan untuk menentukan main tenis, kriteria yang diperhatikan adalah cuaca, angin, dan temperatur. Salah satu atribut merupakan atribut yang menyatakan data solusi per *item* data yang disebut target atribut. Atribut memiliki nilai-nilai yang dinamakan dengan *intance*. Misalkan atribut cuaca mempunyai *intance* berupa cerah, berawan, dan hujan.

### 2.5.2 Jenis-Jenis *Decision tree*

Beberapa metode *decision tree* yang telah dikembangkan saat ini antara lain : algoritma ID3, algoritma C4.5, dan Cart.

### 1) Algoritma ID3

Algoritma ID3 (*Iterative Dichotomiser 3*) pertama kali dikenalkan oleh Quinlan yang digunakan untuk menginduksi *decision tree*. Algoritma ID3 dapat bekerja dengan baik pada semua fitur yang mempunyai tipe data kategorikal (nominal atau ordinal). Dalam perkembangannya, ID3 mengalami perbaikan pada versi berikutnya seperti C4.5 atau C5.0.

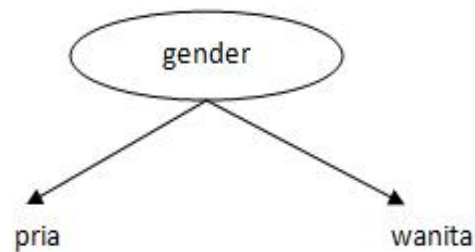
Tabel 2.1 Algoritma Induksi Decision Tree

1.	Dimulai dari node akar.
2.	Untuk semua fitur, hitung nilai entropy untuk semua sampel data pada node.
3.	Pilih fitur dengan interaksi gain yang maksimal.
4.	Gunakan fitur tersebut sebagai node pemecahan menjadi cabang.
5.	Lakukan secara rekursif pada setiap cabang yang dibuat dengan mengulangi langkah 2 sampai 4 hingga semua data dalam setiap node hanya memberikan satu label kelas. Node yang tidak dapat dipecah lagi merupakan daun yang berisi keputusan (label kelas).

### 2) Algoritma C4.5

Algoritma C4.5 digunakan untuk klasifikasi data dengan membentuk pohon keputusan. Algoritma C4.5 diperkenalkan oleh Quinlan sebagai versi perbaikan dari ID3. Dalam ID3, induksi decision tree hanya bisa dilakukan pada fitur bertipe kategorikal (nominal atau ordinal), sedangkan tipe numerik (interval atau rasio) tidak dapat digunakan. Perbaikan yang membedakan algoritma C4.5 dari ID3 adalah dapat menangani fitur dengan tipe numerik, melakukan pemotongan (*pruning*) *decision tree*, dan penurunan (*deriving*) *rule set*. Perhitungan *Decision Tree C4.5* ditunjukkan pada tabel 2.2. Algoritma C4.5 juga menggunakan kriteria gain dalam menentukan fitur yang menjadi pemecah node pada pohon yang di induksi.





Gambar 2.2 Syarat Pengujian Fitur Biner

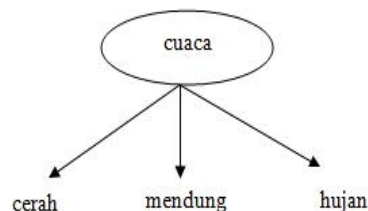
Yang menjadi hal penting dalam induksi *decision tree* adalah bagaimana menyatakan syarat pengujian pada node. Ada 3 kelompok penting dalam syarat penyajian node :

a. Fitur biner

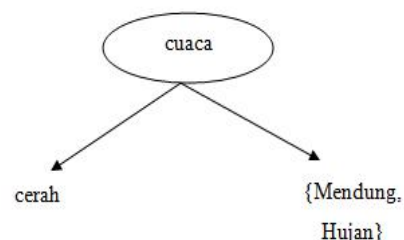
Fitur yang hanya mempunyai dua nilai berbeda disebut dengan fitur biner. Syarat pengujian ketika fitur ini menjadi node (akar maupun internal) hanya punya dua pilihan cabang. Contoh pemecahannya disajikan pada gambar 2.2.

b. Fitur bertipe kategorikal

Untuk fitur yang nilainya bertipe kategorikal (nominal atau ordinal) bisa mempunyai nilai berbeda. Contohnya adalah fitur “cuaca” mempunyai 3 nilai berbeda, dan bisa mempunyai banyak kombinasi syarat pengujian pemecahan. Secara umum ada 2, yaitu pemecahan biner (*binary splitting*) dan (*multi splitting*).



Gambar 2.3 Multi Splitting

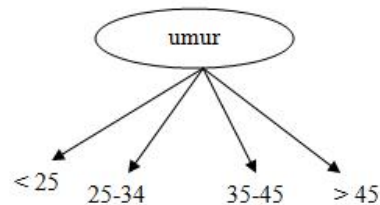


Gambar 2.4 Binary Splitting

c. Fitur bertipe numerik

Untuk fitur bertipe numerik, syarat pengujian dalam node (akar maupun internal) dinyatakan dengan pengujian perbandingan ( $A <$

$v$ ) atau ( $A \geq v$ ) dengan hasil biner, atau untuk multi dengan hasil berupa jangkauan nilai dalam bentuk  $v_i \leq A < v_i + 1$ , untuk  $i=1,2,3,\dots,k$ .

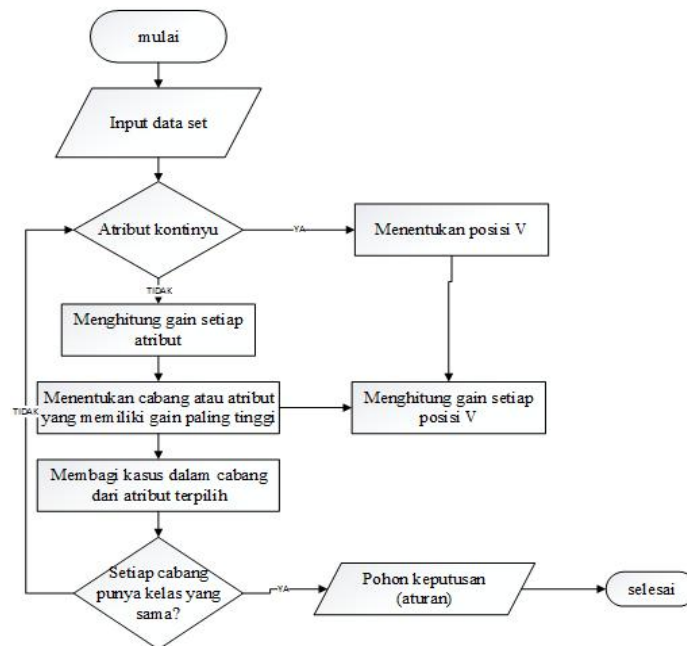


Gambar 2.5 Multi Splitting

Secara umum algoritma C4.5 untuk membangun pohon keputusan adalah sebagai berikut:

- Pilih atribut sebagai node akar.
- Buatlah cabang untuk setiap nilai.
- Bagi kasus dalam cabang.
- Ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama.

Berikut ini akan dijelaskan secara lebih detail algoritma C4.5 menggunakan *flowchart* yang disajikan pada **gambar 2.6**.



Gambar 2.6 *Flowcart* Algoritma *Decision Tree* C4.5

Untuk memilih atribut sebagai simpul akar (*root node*) atau simpul dalam (*internal node*), didasarkan pada nilai *information gain* tertinggi dari atribut-atribut yang ada. Sebelum perhitungan *information gain*, akan dilakukan perhitungan *entropy*. *Entropy* merupakan distribusi probabilitas dalam teori informasi dan diadopsi kedalam algoritma C4.5 untuk mengukur tingkat homogenitas distribusi kelas dari sebuah himpunan data (*data set*). Semakin tinggi tingkat *entropy* dari sebuah data maka semakin homogen distribusi kelas pada data tersebut. Perhitungan *information gain* menggunakan rumus 2.1, sedangkan *entropy* menggunakan rumus 2.2.

$$Gain(S,A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i) \dots\dots(2.1)$$

dimana,

S : Himpunan kasus

A : Atribut

n : Jumlah partisi atribut A

|S<sub>i</sub>|: Jumlah kasus pada partisi ke i

|S| : Jumlah kasus dalam S

$$Entropy(S) = - \sum_{i=1}^n p_i * \log_2 p_i \dots\dots\dots(2.2)$$

dimana,

S : Himpunan kasus

n : Jumlah partisi S

p<sub>i</sub> : Proporsi dari S<sub>i</sub> terhadap S

Selain *Information Gain* kriteria yang lain untuk memilih atribut sebagai pemecah adalah *Rasio Gain*. Perhitungan rasio gain menggunakan rumus 2.3, sedangkan split information menggunakan rumus 2.4.

$$GainRasio(S,A) = \frac{Gain(S,A)}{SplitInformation(S,A)} \dots\dots\dots(2.3)$$

$$SplitInformation(S,A) = - \sum_{i=1}^c \frac{S_i}{S} \log_2 \frac{S_i}{S} \dots\dots\dots(2.4)$$

Dimana  $S_1$  sampai  $S_c$  adalah  $c$  subset yang dihasilkan dari pemecahan  $S$  dengan menggunakan atribut  $A$  yang mempunyai sebanyak  $c$  nilai.

Untuk mengukur nilai akurasi yang didapat dari hasil pengujian menggunakan rumus 2.5. Sedangkan untuk mengukur tingkat kesalahannya menggunakan rumus 2.6.

$$\text{Akurasi} = \frac{\text{jumlah data yang diprediksi benar}}{\text{jumlah prediksi yang dilakukan}} \quad (2.5)$$

$$\text{Laju error} = \frac{\text{jumlah data yang diprediksi salah}}{\text{jumlah prediksi yang dilakukan}} \quad (2.6)$$

## 2.6 Hipertensi

Semakin tinggi tekanan darah risikonya semakin besar. Bahkan seseorang yang memiliki rata-rata tekanan darahnya tinggi mempunyai kecenderungan menderita penyakit jantung lebih besar dibandingkan dengan mereka yang tekanan darahnya rata-rata rendah. Setiap orang mempunyai tekanan darah yang berbeda-beda, tergantung dari usia dan kegiatan sehari-harinya. Pada umumnya semakin bertambah usia seseorang tekanan darahnya cenderung semakin tinggi. Menurut *The Seventh Report of The Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure* (JNC VII) tekanan darah seseorang dapat dikelompokkan sebagai berikut ( dr. Widharto, 2007).

Tabel 2.2 Derajat Hipertensi Untuk Usia 18 Tahun

Tekanan Sistolik (mmHg)	Tekanan Diastolik (mmHg)	Klasifikasi
< 120	80	Normal
120-139	80-89	Prehipertensi
140-159	90-99	Hipertensi tingkat 1
>160	>100	Hipertensi tingakat 2

Menurut WHO dan *British Hypertension* tingkat keparahan tekanan darah tinggi seseorang ditunjukkan pada Tabel 2.2 dibawah ini (dr.Elizabeth Yasmin, 2007).

Tabel 2.3 Tingkat Keparahan Hipertensi Menurut WHO *British Hypertension*

Tingkat Keparahan Tekanan Darah Tinggi		
Kategori	Tekanan Sistolik mmHg	Tekanan Diastolik mmHg
<b>Tekanan darah</b> Optimal Normal Normal tinggi	Di bawah 120	Di bawah 80
	Dibawah 130 130 - 139	Di bawah 85 85 - 89
<b>Hipertensi</b> Ringan (stadium 1) Sedang (stadium 2) Berat (stadium 3)	140 – 159	90 – 99
	160 – 179	100 – 109
	≥ 180	≥ 110
<b>Hipertensi sistolik terisolasi<sup>a</sup></b> Stadium 1 Stadium 2	140 – 159	Di bawah 90
	≥160	Dibawah 90

<sup>a</sup>Hipertensi sistolik terisolasi berarti tekanan diastoliknya normal, namun tekanan sistolik terlalu tinggi. Keadaan ini biasa terjadi pada orang berusia lanjut.

Tabel 2.4 Klasifikasi Dan Tindak Lanjut Pengukuran Tekanan Darah.<sup>1</sup>

Kategori <sup>2</sup>	Tekanan darah sistolik (mmHg)	Tekanan darah diastolik (mmHg)	Tindak lanjut yang dianjurkan
Pilihan	< 120	< 80	
Normal	< 130	< 85	Cek ulang dalam waktu 2 tahun
Normal Tinggi	130 – 139	85 – 90	Cek ulang dalam waktu 2 tahun Cek ulang dalam waktu 1 tahun <sup>3</sup>

Hipertensi <sup>4</sup>			
Derajat 1 (ringan)	140 – 159	90 – 99	Konfirmasi dalam waktu 2 bulan
Derajat 2 (sedang)	160 – 179	100 – 109	Evaluasi atau rujuk dalam waktu bulan
Derajat 3 (berat)	> 180	> 100	Evaluasi atau rujuk dalam waktu 1 minggu

<sup>1</sup>Dari : *The Sixth report of the Joint National Committee on detection, education and treatment of high blood pressure (JNC VI). Arch Intern Med* 1997;157;2413 (Sumber : Dr. Abdul Gofir, 2002).

<sup>2</sup>Ketika tekanan darah sistolik dan diastolik berada pada kategori yang berbeda, kategori yang lebih tinggi hendaknya dipilih untuk mengklasifikasikan tekanan darah individu. Hipertensi sistolik terisolasi didefinisikan sebagai tekanan darah sistolik 140 mmHg atau lebih dan tekanan darah diastolik kurang dari 90 mmHg.

<sup>3</sup>Pertimbangan konseling tentang modifikasi gaya hidup Tabel 2.5.

<sup>4</sup>Individu yang berusia 18 tahun ke atas yang tidak meminum obat anti hipertensi dan tidak menderita sakit akut. Berdasarkan rata-rata dua atau lebih hasil pembacaan pada dua atau lebih kesempatan setelah skrining (*Screening*) awal.

Hipertensi yang berlangsung lama tidak harus memerlukan terapi farmakologis. Pendekatan non farmakologis dan penilaian individual terhadap rasio untung rugi terapi obatnya hendaknya dilakukan sebelum penatalaksanaan farmakologis pada pasien dengan hipertensi derajat 1 (tekanan diastolik < 100 mmHg; tekanan sistoliknya < 160 mmHg). Beberapa pasien yang mempunyai tekanan dalam rentang tersebut tidak memerlukan terapi obat jika mereka tidak terbukti mengalami kerusakan organ target atau penyakit kardiovaskular atau diabetes yang bersamaan (Dr. Abdul Gofir, 2002 ).

### 2.6.1 Terapi Non Farmakologis

Modifikasi gaya hidup dapat mempunyai pengaruh yang mendasar terhadap morbiditas dan mortalitas. Diet yang kaya buah-buahan, sayuran, dan rendah lemak dapat menurunkan tekanan darah. Terapi tambahan dapat mencegah atau mengurangi hipertensi akibat kardiovaskular seperti ditunjukkan pada Tabel 2.5.

Tabel 2.5 Modifikasi Gaya Hidup Untuk Pencegahan Dan Penatalaksanaan Hipertensi.<sup>1</sup>

1	Kurangi berat badan jika berlebih
2	Batasi asupan alkohol, etanol tidak lebih dari 1 oz (30 ml), dan 0.5 oz (15 ml) etanol setiap hari untuk wanita
3	Tingkatkan aktivitas fisik aerobik (30 – 45 menit hampir setiap hari dalam satu minggu)
4	Kurangi asupan natrium tidak lebih dari 100 mmol/ hari (2.4 gram natrium atau 6 gram natrium klorida)
5	Pertahankan asupan kalium yang adekuat dalam diet (kira-kira 90 mmol/hari)
6	Pertahankan intake kalsium dan magnesium yang adekuat dalam diet untuk kesehatan secara umum
7	Berhenti merokok dan kurangi asupan lemak jenuh dalam diet dan kolesterol untuk kesehatan kardiovaskular secara keseluruhan

<sup>1</sup>Dari : *The Sixth report of the Joint National Committee on detection, education and treatment of high blood pressure (JNC VI). Arch Intern Med 1997;157;2413.*

Tekanan sistolik seseorang terus meningkat sampai usia 80 tahun. Sedangkan tekanan diastolik terus meningkat sampai usia 55-60 tahun, kemudian berkurang secara perlahan atau bahkan turundrastis. Hipertensi dapat dibedakan menjadi dua, yaitu hipertensi primer dan hipertensi sekunder. Hipertensi primer merupakan hipertensi yang tidak diketahui penyebabnya. Lebih dari 90% penderita hipertensi termasuk

jenis hipertensi primer. Banyaknya pakar menduga stres merupakan penyebab utama hipertensi primer. Hipertensi sekunder adalah hipertensi yang disebabkan oleh penyakit lain seperti diabetes, dan penyakit pembuluh darah ( dr. Widharto, 2007).

Faktor resiko hipertensi meliputi : kelebihan berat badan, kurang berolahraga, mengkonsumsi makanan berkadar garam tinggi, kurang konsumsi buah dan sayuran segar, terlalu banyak minum alkohol, dan perokok aktif (Dr Elizabeth Yasmine, 2007).

### **2.6.2 Emergensi Hipertensi**

Hipertensi gawat (*hypertensive emergencies*) merupakan situasi dimana tekanan darah harus diturunkan dalam waktu beberapa jam. Ini termasuk pasien dengan hipertensi berat asimtomatik (tekanan darah sistolik > 220 mmHg atau tekanan darah diastolik > 125 mmHg) beresiko pada kematian.

## **2.7 Tekanan Darah**

Tekanan darah adalah gaya ( atau dorongan ) darah ke dinding arteri saat darah dipompa keluar dari jantung keseluruh tubuh. Hasil pengukuran tekanan darah berupa dua angka, yang menunjukkan tekanan sistolik dan diastolik (misalnya : 120/80, disebut “seratus dua puluh per delapan puluh”). Angka yang diatas menunjukkan tekanan sistolik, yaitu tekanan dari arteri saat jantung memompa darah melalui pembuluh tersebut. Angka yang dibawah menunjukkan tekanan diastolik, yaitu tekanan diarteri saat jantung berelaksasi diantara dua denyutan (kontraksi).

## **2.8 Penelitian Terdahulu**

Penelitian terdahulu yang menggunakan metode *Decision tree* C4.5 dalam penelitian berjudul “ *Klasifikasi Data Kejadian Luar Biasa Campak Menggunakan Metode Decission Tree C4.5* ”, dibuat oleh sulistyowati (Teknik Informatika STMIK Palangkaraya). Tujuan dari penelitian ini yaitu untuk mengetahui kinerja algoritma *decision tree* C4.5 dalam



melakukan klasifikasi data KLB suatu penyakit serta untuk mengetahui daerah yang menederita kejadian luar biasa (KLB) atau non KLB berdasarkan data *survailen* penyakit campak disuatu wilayah. Adapun data yang diambil dalam penelitian ini adalah data yang diperoleh dari dinas kesehatan provinsi DIY *survailen* penyakit campak dan jumlah data yang dikumpulkan adalah 648 data yang selanjutnya telah dibagi kedalam 2 bagian yaitu data pembelajaran dan data pengujian. Atribut yang digunakan adalah 14 kecamatan yang digunakan sebagai sampel data penelitian dan masing-masing kecamatan mempunyai 7 buah atribut. Cakupan Vitamin A, cakupan imunisasi, jumlah balita, jumlah bayi, PHBS, penghuni rumah, cakupan gizi. Target kelas berupa status kecamatan KLB atau Non KLB. Pengujian sistem dilakukan dengan pembentukan rule terlebih dahulu dari 500 data training, dan 109 data uji. Diperoleh hasil bahwa metode ini berhasil mengklasifikasikan 92 data yang benar. Dengan demikian diperoleh akurasi keberhasilan sebesar 84,40% dan nilai RSE sebesar 0,352.

Penelitian lain tentang metode *decision tree* C4.5 dilakukan oleh Bambang Hermanto dan Azhari SN dari FMIPA UGM yang berjudul “*Klasifikasi Nilai Kelayakan Calon Debitur Baru Decision tree C4.5*”. Adapun data yang digunakan yaitu data calon nasabah debitur baru sepeda motor dengan keluaran berupa lunas atau tsriksn. Penelitian ini bertujuan untuk memperkirakan calon debitur baru berpotensi kredit lunas. Atribut yang digunakan yaitu jenis pekerjaan, besar penghasilan, tenor, nilai angsuran, dan status kredit debitur. Jumlah data yang gunakan sebanyak 5 *record*. Diperoleh akurasi keberhasilan sebesar 70%.

Penelitian lain tentang metode *decision tree* C4.5 dilakukan oleh Ade Yuliana dan Duwi Bayu Pratomo dari Politeknik TEDC Bandung yang berjudul “*Algoritma Decision tree (C4.5) Untuk Memprediksi Kepuasan Mahasiswa Terhadap Kinerja Dosen Politeknik Tedc Bandung*”. Adapun data yang digunakan yaitu mahasiswa teknik informatika semester ganjil tahun 2015/2016 di Politeknik TEDC Bandung. Atribut yang

digunakan yaitu tangible (bukti langsung), reability (keandalan), responsiveness (daya tanggap), assurance (jaminan), dan empathy (empati). Tujuan dari penelitian ini yaitu memprediksi kepuasan mahasiswa terhadap kinerja dosen. Hasil akurasi yang di dapat yaitu sebesar 94,62%.