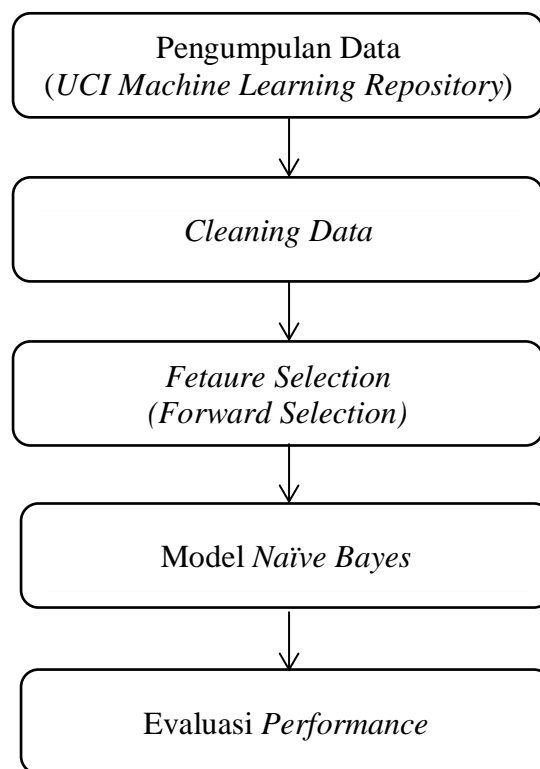


## BAB III

### METODOLOGI PENELITIAN

#### 3.1. Alur Penelitian

Alur penelitian adalah konsep atau gambaran dari penelitian yang akan dilakukan. Penjabaran alur penelitian tersebut dapat dilihat pada gambar diagram 3.1



Gambar 3.1 Diagram Alur Penelitian

#### 3.2. Pengumpulan Data

Berdasarkan diagram diatas, penelitian dimulai dari pengumpulan data dimana data yang digunakan adalah data open source atau data public yang diperoleh dari *UCI Machine Learning Repository* dengan nama *Early stage diabetes risk prediction dataset* yang disumbangkan pada tanggal 12 agustus 2020. Dan terkait informasi pengumpulan datasetnya menggunakan kuesioner langsung dari pasien

Rumah Sakit Sylhet Diabetes di Sylhet, Bangladesh dan disetujui oleh seorang dokter. Berikut adalah link url dari dataset tersebut.

<https://archive.ics.uci.edu/ml/datasets/Early+stAge+diabetes+risk+prediction+data+set.#>. Terdapat 520 data pasien penderita penyakit diabetes dan 17 atribut, dimana satu atribut digunakan sebagai atribut target. Masing-masing dari atribut tersebut dapat dilihat pada tabel 3.1.

Tabel 3.1 Atribut Dataset Penyakit Diabetes.

<i>Age</i>	Usia
<i>Gender</i>	Jenis Kelamin
<i>Polyuria</i>	Sering buang air kecil
<i>polydipsia</i>	Rasa haus yang berlebihan
<i>sudden weight loss</i>	Penurunan berat badan yang signifikan
<i>weakness</i>	Kelelahan/ Lemas
<i>polyphagia</i>	Nafsu makan meningkat
<i>genital thrush</i>	Infeksi jamur kelamin
<i>visual blurring</i>	Pandangan Kabur atau tidak jelas
<i>itching</i>	Gatal dibagian tubuh tertentu dan sulit sembuh
<i>Irritability</i>	Emosi tidak stabil
<i>delayed healing</i>	Luka sulit sembuh
<i>partial paresis</i>	Lumpuh atau merasa lemas
<i>muscle stiffness</i>	Badan tidak seimbang dan merasa kaku
<i>alopecia</i>	Kebotakan/ rambut rontok
<i>obesity</i>	Obisitas
<i>class</i>	<i>Positive</i> atau <i>Negative</i>

### 3.3. *Cleaning Data*

Pada tahap ini dilakukan *cleaning data* pada atribut *Age*, karena berdasarkan informasi dari *UCI machine learning*, *Age* yang digunakan yaitu dari usia dari 20-65 tahun, sedangkan dataset yang ada terdapat *Age* dari usia 16-90 tahun oleh sebab itu dilakukan *cleaning data* menggunakan teknik *filter examples* pada tool rapidminer, sehingga didapat data yang siap digunakan yaitu sebanyak 468 data.

Dengan jumlah *class Positive* sebanyak 284 data dan *class Negative* sebanyak 184 data.

### 3.4. *Feature Selection (Forward Selection)*

Berikut merupakan cara kerja *feature forward selection* dengan mengambil empat atribut yang ada pada dataset diabetes. Atribut tersebut yaitu, *Gender*, *Polyuria*, *Polydipsia*, *Sudden Weight Loss*.

- Iterasi 1

Masing-masing fitur atau atribut dilakukan evaluasi performance untuk melihat akurasi tertinggi. Atribut tertinggi akan dijadikan acuan untuk dilakukan iterasi ke 2.

- *Gender* : 69.03%
- *Polyuria* : 84.19%
- *Polydipsia* : 80.14%
- *Sudden Weight Loss* : 69.02%

- Iterasi 2

Pada iterasi 2 atribut *Polyuria* akan dijadikan acuan yang akan ditambah dengan atribut lainnya.

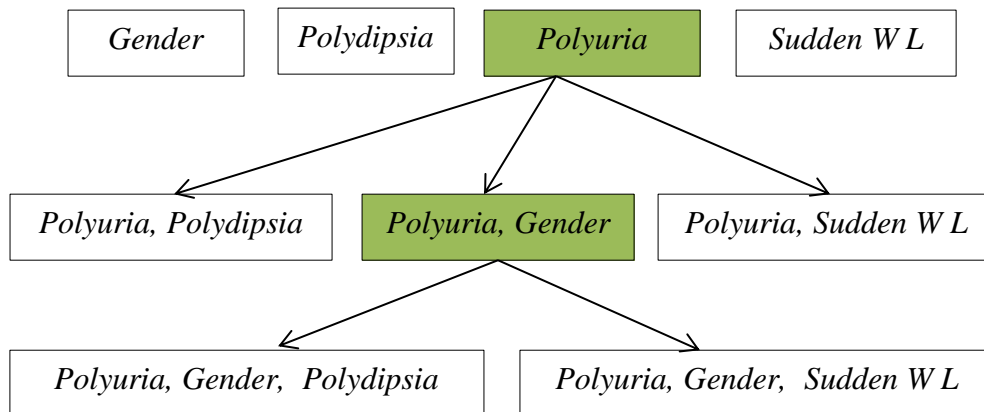
- *Polyuria, Gender* : 89.11%
- *Polyuria, Polydipsia* : 88.03 %
- *Polyuria, Sudden Weight Loss* : 84.19 %

Pada iterasi 2 diperoleh akurasi tertinggi pada atribut *Polyuria* dan *Gender*. Kedua atribut ini digunakan sebagai acuan untuk dilakukan iterasi 3.

- Iterasi 3

Pada iterasi 3 atribut *Polyuria* dan *Gender* akan dijadikan acuan yang akan ditambah dengan atribut lainnya.

- *Polyuria, Gender, Polydipsia* : 89.09 %
- *Polyuria, Gender, Sudden Weight Loss* : 89.09 %



Gambar 3.2 Cara Kerja *Feature Forward Selection*

Dari gambar 3.2 diatas hasil performance menggunakan *feature forward selection* dengan mempertahankan dua atribut yang berpengaruh yaitu *Polyuria* dan *Gender*.

**3.5. Menghitung Nilai Probabilitas Masing-Masing Atribut**

untuk menghitung nilai probabilitas dari masing-masing atribut menggunakan rumus algoritma *naïve bayes* di bawah ini.

$$P(x|y) = \frac{P(y|x)P(x)}{P(y)} \dots\dots\dots(1)$$

$$P(y) = \sum_{n=1}^n P(y|x)P(x) \dots\dots\dots(2)$$

**1. Menghitung Probabilitas Atribut Age**

Pada atribut *Age* (usia) berisikan data pasien diabetes mulai dari usia 20-65, pada penelitian ini dilakukan pengelompokkan rentang usia dengan keterangan usia 20-45 dikategorikan *Masture* sedangkan usia 45-65 dikategorikan *Old*.

a) Menghitung Probabilitas *Age Masture (Positive)*

$$\frac{Jumlah\ Masture\ Postive}{Jumlah\ Class\ Postive} = \frac{127}{284} = 0,44718309$$

b) Menghitung Probabilitas *Age Masture (Negative)*

$$\frac{Jumlah\ Masture\ Negative}{Jumlah\ Class\ Negative} = \frac{103}{184} = 0,559782609$$

c) Menghitung Probabilitas *Age Old (Positive)*

$$\frac{\text{Jumlah Old Postive}}{\text{Jumlah Class Postive}} = \frac{157}{284} = 0,552816901$$

d) Menghitung Probabilitas *Age Old (Negative)*

$$\frac{\text{Jumlah Old Negative}}{\text{Jumlah Class Negative}} = \frac{157}{184} = 0,440217391$$

Tabel 3.2 Nilai Probabilitas *Age*

<i>Age</i>	<i>Positive</i>	<i>Negative</i>
<i>Masture</i>	0,447183099	0,559782609
<i>Old</i>	0,552816901	0,440217391

## 2. Menghitung Probabilitas Atribut *Gender*

Pada atribut gender terdapat value male dan female. Berikut adalah perhitungan dari masing-masing value male dan female pada atribut gender.

a) Menghitung Probabilitas *Gender Male (Positive)*

$$\frac{\text{Jumlah Male Postive}}{\text{Jumlah Class Postive}} = \frac{126}{284} = 0,443661972$$

b) Menghitung Probabilitas *Gender Male (Negative)*

$$\frac{\text{Jumlah Male Negative}}{\text{Jumlah Class Negative}} = \frac{165}{184} = 0,89673913$$

c) Menghitung Probabilitas *Gender Female (Positive)*

$$\frac{\text{Jumlah Female Positive}}{\text{Jumlah Class Positive}} = \frac{158}{284} = 0,556338028$$

d) Menghitung Probabilitas *Gender Female (Negative)*

$$\frac{\text{Jumlah Male Negative}}{\text{Jumlah Class Negative}} = \frac{19}{184} = 0,10326087$$

Tabel 3.3 Nilai Probabilitas Atribut *Gender*

<i>Gender</i>	<i>Positive</i>	<i>Negative</i>
<i>Male</i>	0,443661972	0,89673913
<i>Female</i>	0,556338028	0,10326087

### 3. Menghitung Probabilitas Atribut *Polyuria*

Pada atribut *polyuria* terdapat value *Yes* dan *No*. Berikut adalah perhitungan dari masing-masing value *Yes* dan *No* pada atribut *polyuria* .

- a) Menghitung Probabilitas *Polyuria Yes (Positive)*

$$\frac{\text{Jumlah Yes Postive}}{\text{Jumlah Class Postive}} = \frac{215}{284} = 0,757042254$$

- b) Menghitung Probabilitas *Polyuria Yes (Negative)*

$$\frac{\text{Jumlah Yes Negative}}{\text{Jumlah Class Negative}} = \frac{5}{184} = 0,027173913$$

- c) Menghitung Probabilitas *Polyuria No (Positive)*

$$\frac{\text{Jumlah No Positive}}{\text{Jumlah Class Positive}} = \frac{69}{284} = 0,242957746$$

- d) Menghitung Probabilitas *Polyuria No (Negative)*

$$\frac{\text{Jumlah No Negative}}{\text{Jumlah Class Negative}} = \frac{179}{184} = 0,972826087$$

Tabel 3.4 Nilai Probabilitas *Polyuria*

<i>Polyuria</i>	<i>Positive</i>	<i>Negative</i>
<i>Yes</i>	0,757042254	0,027173913
<i>No</i>	0,242957746	0,972826087

### 4. Menghitung Probabilitas Atribut *Polydipsia*

Pada atribut *polydipsia* terdapat value *Yes* dan *No*. Berikut adalah perhitungan dari masing-masing value *Yes* dan *No* pada atribut *polydipsia* .

- a) Menghitung Probabilitas *Polydipsia Yes (Positive)*

$$\frac{\text{Jumlah Yes Postive}}{\text{Jumlah Class Postive}} = \frac{199}{284} = 0,700704225$$

- b) Menghitung Probabilitas *Polydipsia Yes (Negative)*

$$\frac{\text{Jumlah Yes Negative}}{\text{Jumlah Class Negative}} = \frac{8}{184} = 0,043478261$$

- c) Menghitung Probabilitas *Polydipsia No (Positive)*

$$\frac{\text{Jumlah No Positive}}{\text{Jumlah Class Positive}} = \frac{85}{176} = 0,299295775$$

- d) Menghitung Probabilitas *Polydipsia No (Negative)*

$$\frac{\text{Jumlah No Negative}}{\text{Jumlah Class Negative}} = \frac{176}{184} = 0,956521739$$

Tabel 3.5 Nilai Probabilitas *Polydipsia*

<i>Polydipsia</i>	<i>Positive</i>	<i>Negative</i>
<i>Yes</i>	0,700704225	0,043478261
<i>No</i>	0,299295775	0,956521739

### 5. Menghitung Probabilitas Atribut *Sudden Weight Loss*

Pada atribut *sudden weight loss* terdapat value *Yes* dan *No*. Berikut adalah perhitungan dari masing-masing value *Yes* dan *No* pada atribut *sudden weight loss*.

- a) Menghitung Probabilitas *Sudden Weight Loss Yes (Positive)*

$$\frac{\text{Jumlah Yes Postive}}{\text{Jumlah Class Postive}} = \frac{168}{284} = 0,591549296$$

- b) Menghitung Probabilitas *Sudden Weight Loss Yes (Negative)*

$$\frac{\text{Jumlah Yes Negative}}{\text{Jumlah Class Negative}} = \frac{29}{184} = 0,157608696$$

- c) Menghitung Probabilitas *Sudden Weight Loss No (Positive)*

$$\frac{\text{Jumlah No Positive}}{\text{Jumlah Class Positive}} = \frac{116}{176} = 0,408450704$$

- d) Menghitung Probabilitas *Sudden Weight Loss No (Negative)*

$$\frac{\text{Jumlah No Negative}}{\text{Jumlah Class Negative}} = \frac{155}{184} = 0,842391304$$

Tabel 3.6 Nilai Probabilitas *Sudden Weight Loss*

<i>sudden weight loss</i>	<i>Positive</i>	<i>Negative</i>
<i>Yes</i>	0,591549296	0,157608696
<i>No</i>	0,408450704	0,842391304

### 6. Menghitung Probabilitas Atribut *Weakness*

Pada atribut *weakness* terdapat value *Yes* dan *No*. Berikut adalah perhitungan dari masing-masing value *Yes* dan *No* pada atribut *weakness*.

- a) Menghitung Probabilitas *Weakness Yes (Positive)*

$$\frac{\text{Jumlah Yes Postive}}{\text{Jumlah Class Postive}} = \frac{190}{284} = 0,669014085$$

b) Menghitung Probabilitas *Weakness Yes (Negative)*

$$\frac{\text{Jumlah Yes Negative}}{\text{Jumlah Class Negative}} = \frac{80}{184} = 0,434782609$$

c) Menghitung Probabilitas *Weakness No (Positive)*

$$\frac{\text{Jumlah No Positive}}{\text{Jumlah Class Positive}} = \frac{94}{176} = 0,330985915$$

d) Menghitung Probabilitas *Weakness No (Negative)*

$$\frac{\text{Jumlah No Negative}}{\text{Jumlah Class Negative}} = \frac{104}{184} = 0,565217391$$

Tabel 3.7 Nilai Probabilitas *Weakness*

<i>weakness</i>	<i>Positive</i>	<i>Negative</i>
<i>Yes</i>	0,669014085	0,434782609
<i>No</i>	0,330985915	0,565217391

## 7. Menghitung Probabilitas Atribut *Polyphagia*

Pada atribut *polyphagia* terdapat value *Yes* dan *No*. Berikut adalah perhitungan dari masing-masing value *Yes* dan *No* pada atribut *polyphagia*.

a) Menghitung Probabilitas *Polyphagia Yes (Positive)*

$$\frac{\text{Jumlah Yes Postive}}{\text{Jumlah Class Postive}} = \frac{162}{284} = 0,570422535$$

b) Menghitung Probabilitas *Polyphagia Yes (Negative)*

$$\frac{\text{Jumlah Yes Negative}}{\text{Jumlah Class Negative}} = \frac{34}{184} = 0,184782609$$

c) Menghitung Probabilitas *Polyphagia No (Positive)*

$$\frac{\text{Jumlah No Positive}}{\text{Jumlah Class Positive}} = \frac{122}{176} = 0,429577465$$

d) Menghitung Probabilitas *Polyphagia No (Negative)*

$$\frac{\text{Jumlah No Negative}}{\text{Jumlah Class Negative}} = \frac{150}{184} = 0,815217391$$

Tabel 3.8 Nilai Probabilitas *Polyphagia*

<i>Polyphagia</i>	<i>Positive</i>	<i>Negative</i>
<i>Yes</i>	0,570422535	0,184782609
<i>No</i>	0,429577465	0,815217391



### 8. Menghitung Probabilitas Atribut *Genital Thrush*

Pada atribut *genital thrush* terdapat value *Yes* dan *No*. Berikut adalah perhitungan dari masing-masing value *Yes* dan *No* pada atribut *genital thrush*.

- a) Menghitung Probabilitas *Genital Thrush Yes (Positive)*

$$\frac{\text{Jumlah Yes Postive}}{\text{Jumlah Class Postive}} = \frac{68}{284} = 0,23943662$$

- b) Menghitung Probabilitas *Genital Thrush Yes (Negative)*

$$\frac{\text{Jumlah Yes Negative}}{\text{Jumlah Class Negative}} = \frac{33}{184} = 0,179347826$$

- c) Menghitung Probabilitas *Genital Thrush No (Positive)*

$$\frac{\text{Jumlah No Positive}}{\text{Jumlah Class Positive}} = \frac{216}{176} = 0,76056338$$

- d) Menghitung Probabilitas *Genital Thrush No (Negative)*

$$\frac{\text{Jumlah No Negative}}{\text{Jumlah Class Negative}} = \frac{151}{184} = 0,820652174$$

Tabel 3.9 Nilai Probabilitas *Genital Thrush*

Genital thrush	<i>Positive</i>	<i>Negative</i>
<i>Yes</i>	0,23943662	0,179347826
<i>No</i>	0,76056338	0,820652174

### 9. Menghitung Probabilitas Atribut *Visual blurring*

Pada atribut *visual blurring* terdapat value *Yes* dan *No*. Berikut adalah perhitungan dari masing-masing value *Yes* dan *No* pada atribut *visual blurring*.

- a) Menghitung Probabilitas *Visual Blurring Yes (Positive)*

$$\frac{\text{Jumlah Yes Postive}}{\text{Jumlah Class Postive}} = \frac{148}{284} = 0,521126761$$

- b) Menghitung Probabilitas *Visual Blurring Yes (Negative)*

$$\frac{\text{Jumlah Yes Negative}}{\text{Jumlah Class Negative}} = \frac{43}{184} = 0,233695652$$

- c) Menghitung Probabilitas *Visual Blurring No (Positive)*

$$\frac{\text{Jumlah No Positive}}{\text{Jumlah Class Positive}} = \frac{136}{176} = 0,766304348$$

d) Menghitung Probabilitas *Visual Blurring No (Negative)*

$$\frac{\text{Jumlah No Negative}}{\text{Jumlah Class Negative}} = \frac{141}{184} = 0,820652174$$

Tabel 3.10 Nilai Probabilitas *Visual Blurring*

<i>Visual blurring</i>	<i>Positive</i>	<i>Negative</i>
<i>Yes</i>	0,521126761	0,233695652
<i>No</i>	0,478873239	0,766304348

### 10. Menghitung Probabilitas Atribut *Itching*

Pada atribut *itching* terdapat value *Yes* dan *No*. Berikut adalah perhitungan dari masing-masing value *Yes* dan *No* pada atribut *itching*.

a) Menghitung Probabilitas *Itching Yes (Positive)*

$$\frac{\text{Jumlah Yes Postive}}{\text{Jumlah Class Postive}} = \frac{130}{284} = 0,457746479$$

b) Menghitung Probabilitas *Itching Yes (Negative)*

$$\frac{\text{Jumlah Yes Negative}}{\text{Jumlah Class Negative}} = \frac{84}{184} = 0,456521739$$

c) Menghitung Probabilitas *Itching No (Positive)*

$$\frac{\text{Jumlah No Positive}}{\text{Jumlah Class Positive}} = \frac{154}{176} = 0,542253521$$

d) Menghitung Probabilitas *Itching No (Negative)*

$$\frac{\text{Jumlah No Negative}}{\text{Jumlah Class Negative}} = \frac{100}{184} = 0,543478261$$

Tabel 3.11 Nilai Probabilitas *Itching*

<i>Itching</i>	<i>Positive</i>	<i>Negative</i>
<i>Yes</i>	0,457746479	0,456521739
<i>No</i>	0,542253521	0,543478261

### 11. Menghitung Probabilitas Atribut *Irritability*

Pada atribut *irritability* terdapat value *Yes* dan *No*. Berikut adalah perhitungan dari masing-masing value *Yes* dan *No* pada atribut *irritability*.

a) Menghitung Probabilitas *Irritability Yes (Positive)*

$$\frac{\text{Jumlah Yes Postive}}{\text{Jumlah Class Postive}} = \frac{91}{284} = 0,320422535$$

- b) Menghitung Probabilitas *Irritability Yes (Negative)*

$$\frac{\text{Jumlah Yes Negative}}{\text{Jumlah Class Negative}} = \frac{13}{184} = 0,070652174$$

- c) Menghitung Probabilitas *Irritability No (Positive)*

$$\frac{\text{Jumlah No Positive}}{\text{Jumlah Class Positive}} = \frac{193}{176} = 0,679577465$$

- d) Menghitung Probabilitas *Irritability No (Negative)*

$$\frac{\text{Jumlah No Negative}}{\text{Jumlah Class Negative}} = \frac{171}{184} = 0,929347826$$

Tabel 3.12 Nilai Probabilitas *Irritability*

<i>Irritability</i>	<i>Positive</i>	<i>Negative</i>
<i>Yes</i>	0,320422535	0,070652174
<i>No</i>	0,679577465	0,929347826

## 12. Menghitung Probabilitas Atribut *Delayed Healing*

Pada atribut *delayed healing* terdapat value *Yes* dan *No*. Berikut adalah perhitungan dari masing-masing value *Yes* dan *No* pada atribut *delayed healing*.

- a) Menghitung Probabilitas *Delayed Healing Yes (Positive)*

$$\frac{\text{Jumlah Yes Postive}}{\text{Jumlah Class Postive}} = \frac{133}{284} = 0,468309859$$

- b) Menghitung Probabilitas *Delayed Healing Yes (Negative)*

$$\frac{\text{Jumlah Yes Negative}}{\text{Jumlah Class Negative}} = \frac{71}{184} = 0,385869565$$

- c) Menghitung Probabilitas *Delayed Healing No (Positive)*

$$\frac{\text{Jumlah No Positive}}{\text{Jumlah Class Positive}} = \frac{151}{176} = 0,531690141$$

- d) Menghitung Probabilitas *Delayed Healing No (Negative)*

$$\frac{\text{Jumlah No Negative}}{\text{Jumlah Class Negative}} = \frac{113}{184} = 0,614130435$$

Tabel 3.13 Nilai Probabilitas *Delayed Healing*

<i>Delayed Healing</i>	<i>Positive</i>	<i>Negative</i>
<i>Yes</i>	0,468309859	0,385869565
<i>No</i>	0,531690141	0,614130435

### 13. Menghitung Probabilitas Atribut *Partial Paresis*

Pada atribut *partial paresis* terdapat value *Yes* dan *No*. Berikut adalah perhitungan dari masing-masing value *Yes* dan *No* pada atribut *partial paresis*.

- a) Menghitung Probabilitas *Partial Paresis Yes (Positive)*

$$\frac{\text{Jumlah Yes Postive}}{\text{Jumlah Class Postive}} = \frac{167}{284} = 0,588028169$$

- b) Menghitung Probabilitas *Partial Paresis Yes (Negative)*

$$\frac{\text{Jumlah Yes Negative}}{\text{Jumlah Class Negative}} = \frac{24}{184} = 0,130434783$$

- c) Menghitung Probabilitas *Partial Paresis No (Positive)*

$$\frac{\text{Jumlah No Positive}}{\text{Jumlah Class Positive}} = \frac{117}{176} = 0,411971831$$

- d) Menghitung Probabilitas *Partial Paresis No (Negative)*

$$\frac{\text{Jumlah No Negative}}{\text{Jumlah Class Negative}} = \frac{160}{184} = 0,869565217$$

Tabel 3.14 Nilai Probabilitas *Partial Paresis*

<i>Partial Paresis</i>	<i>Positive</i>	<i>Negative</i>
<i>Yes</i>	0,588028169	0,130434783
<i>No</i>	0,411971831	0,869565217

### 14. Menghitung Probabilitas Atribut *Muscle Stiffnes*

Pada atribut *muscle stiffness* terdapat value *Yes* dan *No*. Berikut adalah perhitungan dari masing-masing value *Yes* dan *No* pada atribut *muscle stiffness*.

- a) Menghitung Probabilitas *Muscle Stiffness Yes (Positive)*

$$\frac{\text{Jumlah Yes Postive}}{\text{Jumlah Class Postive}} = \frac{115}{284} = 0,404929577$$

- b) Menghitung Probabilitas *Muscle Stiffness Yes (Negative)*

$$\frac{\text{Jumlah Yes Negative}}{\text{Jumlah Class Negative}} = \frac{46}{184} = 0,25$$

- c) Menghitung Probabilitas *Muscle Stiffness No (Positive)*

$$\frac{\text{Jumlah No Positive}}{\text{Jumlah Class Positive}} = \frac{169}{176} = 0,778169014$$

d) Menghitung Probabilitas *Muscle Stiffness No (Negative)*

$$\frac{\text{Jumlah No Negative}}{\text{Jumlah Class Negative}} = \frac{138}{184} = 0,75$$

Tabel 3.15 Nilai Probabilitas *Muscle Stiffness*

<i>Muscle Stiffness</i>	<i>Positive</i>	<i>Negative</i>
<i>Yes</i>	0,404929577	0,25
<i>No</i>	0,595070423	0,75

### 15. Menghitung Probabilitas Atribut *Alopecia*

Pada atribut *alopecia* terdapat value *Yes* dan *No*. Berikut adalah perhitungan dari masing-masing value *Yes* dan *No* pada atribut *alopecia*.

a) Menghitung Probabilitas *Alopecia Yes (Positive)*

$$\frac{\text{Jumlah Yes Postive}}{\text{Jumlah Class Postive}} = \frac{63}{284} = 0,221830986$$

b) Menghitung Probabilitas *Alopecia Yes (Negative)*

$$\frac{\text{Jumlah Yes Negative}}{\text{Jumlah Class Negative}} = \frac{86}{184} = 0,467391304$$

c) Menghitung Probabilitas *Alopecia No (Positive)*

$$\frac{\text{Jumlah No Positive}}{\text{Jumlah Class Positive}} = \frac{221}{176} = 0,778169014$$

d) Menghitung Probabilitas *Alopecia No (Negative)*

$$\frac{\text{Jumlah No Negative}}{\text{Jumlah Class Negative}} = \frac{98}{184} = 0,532608696$$

Tabel 3.16 Nilai Probabilitas *Alopecia*

<i>Alopecia</i>	<i>Positive</i>	<i>Negative</i>
<i>Yes</i>	0,221830986	0,467391304
<i>No</i>	0,778169014	0,532608696

### 16. Menghitung Probabilitas Atribut *Obesity*

Pada atribut *obesity* terdapat value *Yes* dan *No*. Berikut adalah perhitungan dari masing-masing value *Yes* dan *No* pada atribut *obesity*.

a) Menghitung Probabilitas *Obesity Yes (Positive)*

$$\frac{\text{Jumlah Yes Postive}}{\text{Jumlah Class Postive}} = \frac{55}{284} = 0,19$$

b) Menghitung Probabilitas *Obesity Yes (Negative)*

$$\frac{\text{Jumlah Yes Negative}}{\text{Jumlah Class Negative}} = \frac{24}{184} = 0,13$$

c) Menghitung Probabilitas *Obesity No (Positive)*

$$\frac{\text{Jumlah No Positive}}{\text{Jumlah Class Positive}} = \frac{229}{176} = 0,81$$

d) Menghitung Probabilitas *Obesity No (Negative)*

$$\frac{\text{Jumlah No Negative}}{\text{Jumlah Class Negative}} = \frac{160}{184} = 0,87$$

Tabel 3.17 Nilai Probabilitas *Obesity*

<i>Obesity</i>	<i>Positive</i>	<i>Negative</i>
<i>Yes</i>	0,19	0,13
<i>No</i>	0,81	0,87

### 17. Menghitung Probabilitas Atribut *Class*

Pada atribut *Class* terdapat value *Positive* dan *Negative*. Berikut adalah perhitungan dari masing-masing value *Positive* dan *Negative* pada atribut *Class*.

e) Menghitung Probabilitas *Class Positive*

$$\frac{\text{Jumlah Class Postive}}{\text{Jumlah Total Class}} = \frac{284}{468} = 0,61$$

f) Menghitung Probabilitas *Class Negative*

$$\frac{\text{Jumlah Class Negative}}{\text{Jumlah Total Class}} = \frac{184}{468} = 0,39$$

Tabel 3.18 Nilai Probabilitas *Class*

<i>Class</i>	Nilai
<i>Positive</i>	0,61
<i>Negative</i>	0,39

### 3.5.1. Menghitung *Confusion Matrix*

Berikut adalah perhitungan confusion matrix dengan mengambil data sebanyak 5 record yang akan dijadikan testing untuk mencari nilai akurasi, nilai precision, dan nilai recall.

Tabel 3.19 Data *Record* Yang Akan Dijadikan *Testing*

<b>Atribut</b>	<b>Kasus No 1</b>	<b>Kasus No 2</b>	<b>Kasus No 3</b>	<b>Kasus No 4</b>	<b>Kasus No 5</b>
Age	Masture	Old	Old	Masture	Masture
Gender	Male	Female	Female	Male	Male
Polyuria	Yes	Yes	Yes	No	No
Polydipsia	No	No	No	Yes	No
sudden weight loss	No	No	Yes	Yes	Yes
weakness	No	No	No	Yes	No
Polyphagia	Yes	Yes	Yes	No	No
Genital thrush	No	No	No	No	No
visual blurring	Yes	No	No	Yes	No
Itching	No	No	No	Yes	No
Irritability	No	Yes	Yes	No	No
delayed healing	No	No	No	No	No
partial paresis	Yes	No	No	Yes	No
muscle stiffness	No	No	Yes	Yes	No
Alopecia	Yes	Yes	No	No	No
Obesity	No	No	Yes	No	No
<b>class</b>	<i>Positive</i>	<i>Positive</i>	<i>Positive</i>	<i>Negative</i>	<i>Negative</i>

### 1. Mencari Prediksi Class pada Kasus No 1 (Satu)

Tabel 3.20 Menghitung Probabilitas *Positive* Dan *Negative* Kasus No 1

Atribut	Kasus No 1	<i>Positive</i>	<i>Negative</i>
Age	Masture	0,4471831	0,559782609
Gender	Male	0,44366197	0,89673913
Polyuria	Yes	0,75704225	0,027173913
Polydipsia	No	0,29929577	0,956521739
sudden weight loss	No	0,4084507	0,842391304
weakness	No	0,33098592	0,842391304
Polyphagia	Yes	0,57042254	0,184782609
Genital thrush	No	0,76056338	0,820652174
visual blurring	Yes	0,52112676	0,233695652
Itching	No	0,54225352	0,543478261
Irritability	No	0,67957746	0,929347826
delayed healing	No	0,53169014	0,614130435
partial paresis	Yes	0,58802817	0,130434783
muscle stiffness	No	0,59507042	0,75
Alopecia	Yes	0,22183099	0,467391304
Obesity	No	0,81	0,87
<b>Nilai Probabilitas</b>		0,000016850	0,000004047

Dari tabel diatas nilai probabilitas *positive* pada kasus no 1 sebesar 0,000016850 dan nilai probabilitas *Negative* sebesar 0,000004047. Dari kedua nilai tersebut, nilai probabilitas *positive* lebih besar dari pada nilai probabilitas *Negative*, sehingga prediksi yang dihasilkan sesuai dengan observed class.



## 2. Mencari Prediksi Class pada Kasus No 2 (Dua)

Tabel 3.21 Menghitung Probabilitas *Positive* Dan *Negative* Kasus No 2

Atribut	Kasus No 2	<i>Positive</i>	<i>Negative</i>
Age	Old	0,552816901	0,440217391
Gender	Female	0,556338028	0,10326087
Polyuria	Yes	0,757042254	0,027173913
Polydipsia	No	0,299295775	0,956521739
sudden weight loss	No	0,408450704	0,842391304
weakness	No	0,330985915	0,565217391
Polyphagia	Yes	0,570422535	0,184782609
Genital thrush	No	0,76056338	0,820652174
visual blurring	No	0,478873239	0,766304348
Itching	No	0,542253521	0,543478261
Irritability	Yes	0,320422535	0,070652174
delayed healing	No	0,531690141	0,614130435
partial paresis	No	0,411971831	0,869565217
muscle stiffness	No	0,595070423	0,75
Alopecia	Yes	0,221830986	0,467391304
Obesity	No	0,81	0,87
<b>Nilai Probabilitas</b>		0,000007929	0,000000409

Dari tabel diatas nilai probabilitas *positive* pada kasus no 2 sebesar 0,000007929 dan nilai probabilitas *Negative* sebesar 0,000000409. Dari kedua nilai tersebut, nilai probabilitas *positive* lebih besar dari pada nilai probabilitas *Negative*, sehingga prediksi yang dihasilkan sesuai dengan observed class.

### 3. Mencari Prediksi Class pada Kasus No 3 (Tiga)

Tabel 3.22 Menghitung Probabilitas *Positive* Dan *Negative* Kasus No 3

Atribut	Kasus No 3	<i>Positive</i>	<i>Negative</i>
Age	Old	0,552816901	0,440217391
Gender	Female	0,556338028	0,10326087
Polyuria	Yes	0,757042254	0,027173913
Polydipsia	No	0,299295775	0,956521739
sudden weight loss	Yes	0,591549296	0,157608696
weakness	No	0,330985915	0,565217391
Polyphagia	Yes	0,570422535	0,184782609
Genital thrush	No	0,76056338	0,820652174
visual blurring	No	0,76056338	0,820652174
Itching	No	0,542253521	0,543478261
Irritability	Yes	0,320422535	0,070652174
delayed healing	No	0,531690141	0,614130435
partial paresis	No	0,411971831	0,869565217
muscle stiffness	Yes	0,404929577	0,25
Alopecia	No	0,778169014	0,532608696
Obesity	Yes	0,19	0,13
<b>Nilai Probabilitas</b>		0,000010456	0,000000005

Dari tabel diatas nilai probabilitas *positive* pada kasus no 3 sebesar 0,000010456 dan nilai probabilitas *Negative* sebesar 0,000000005. Dari kedua nilai tersebut, nilai probabilitas *positive* lebih besar dari pada nilai probabilitas *Negative*, sehingga prediksi yang dihasilkan sesuai dengan observed class.

#### 4. Mencari Prediksi Class pada Kasus No 4 (Empat)

Tabel 3.23 Menghitung Probabilitas *Positive* Dan *Negative* Kasus No 4

Atribut	Kasus No 4	<i>Positive</i>	<i>Negative</i>
Age	Masture	0,447183099	0,559782609
Gender	Male	0,443661972	0,89673913
Polyuria	No	0,242957746	0,972826087
Polydipsia	Yes	0,700704225	0,043478261
sudden weight loss	Yes	0,591549296	0,157608696
weakness	Yes	0,669014085	0,434782609
Polyphagia	No	0,429577465	0,815217391
Genital thrush	No	0,76056338	0,820652174
visual blurring	Yes	0,521126761	0,233695652
Itching	Yes	0,457746479	0,456521739
Irritability	No	0,679577465	0,929347826
delayed healing	No	0,531690141	0,614130435
partial paresis	Yes	0,588028169	0,130434783
muscle stiffness	Yes	0,404929577	0,25
Alopecia	No	0,778169014	0,532608696
Obesity	No	0,81	0,87
<b>Nilai Probabilitas</b>		0,000056239	0,000000895

Dari tabel diatas nilai probabilitas *positive* pada kasus no 4 sebesar 0,000056239 dan nilai probabilitas *Negative* sebesar 0,000000895. Dari kedua nilai tersebut, nilai probabilitas *positive* lebih besar dari pada nilai probabilitas *Negative*, sehingga prediksi yang dihasilkan tidak sesuai dengan *observed class*.

## 5. Mencari Prediksi Class pada Kasus No 5 (Lima)

Tabel 3.24 Menghitung Probabilitas *Positive* Dan *Negative* Kasus No 5

Atribut	Kasus No 5	<i>Positive</i>	<i>Negative</i>
Age	Masture	0,447183099	0,55978261
Gender	Male	0,443661972	0,89673913
Polyuria	No	0,242957746	0,97282609
Polydipsia	No	0,299295775	0,95652174
sudden weight loss	Yes	0,591549296	0,1576087
weakness	No	0,330985915	0,56521739
Polyphagia	No	0,429577465	0,81521739
Genital thrush	No	0,76056338	0,82065217
visual blurring	No	0,478873239	0,76630435
Itching	No	0,542253521	0,54347826
Irritability	No	0,679577465	0,92934783
delayed healing	No	0,531690141	0,61413043
partial paresis	No	0,411971831	0,86956522
muscle stiffness	No	0,595070423	0,75
Alopecia	No	0,778169014	0,5326087
Obesity	No	0,81	0,87
<b>Nilai Probabilitas</b>		0,000013320	0,001998665

Dari tabel diatas nilai probabilitas *positive* pada kasus no 4 sebesar 0,000013320 nilai probabilitas *Negative* sebesar 0,001998665. Dari kedua nilai tersebut, nilai probabilitas *Negative* lebih besar dari pada nilai probabilitas *positive* , sehingga prediksi yang dihasilkan sesuai dengan *observed class*.

Dari hasil prediksi ke 5 record tersebut dapat ditampilkan dalam bentuk *confusion matrix* pada tabel di bawah ini.

Tabel 3.25 *Confusion Matrix*

Observed Class	Predicted Class	
	<i>Positive</i>	<i>Negative</i>
<i>Positive</i>	3	0
<i>Negative</i>	1	1

**a. Menghitung Akurasi**

Berikut adalah cara menghitung nilai akurasi dari tabel confusion matrix 3.25 yang diperoleh.

$$\text{Akurasi} = \frac{(3+1)}{(3+1+1+0)} = \frac{4}{5} = 0,8 \times 100\% = 80\%$$

**b. Menghitung Precision**

Berikut adalah cara menghitung nilai precision dari tabel confusion matrix 3.25 yang diperoleh.

$$\text{Precision} = \frac{3}{(3+1)} = \frac{3}{4} = 0,75$$

**c. Menghitung Recall**

Berikut adalah cara menghitung nilai recall dari tabel confusion matrix 3.25 yang diperoleh.

$$\text{Recall} = \frac{3}{(3+0)} = \frac{3}{3} = 1$$