

BAB V

HASIL ANALISIS DAN VISUALISASI

1.1 PERSIAPAN DATA

Tahap ini melakukan beberapa persiapan proses data. Persiapan proses data tersebut yaitu:

1.1.1. Data Awal

Data awal merupakan semua data hasil dari observasi pada Puskesmas sebelum dilakukan penyeleksian data. Data awal dapat dilihat pada table 5.1 :

Tabel 5.1 Data Awal

No	Nama Atribut	Tipe Data
1	Nama	<i>Character</i>
2	Umur	<i>Numeric</i>
3	Jenis Kelamin	<i>Character</i>
4	Alamat	<i>Character</i>
5	Diagnosa Penyakit	<i>Charcter</i>
6	Jenis Malaria	<i>Charcter</i>

Pada tabel 5.1 terdapat atribut seperti nama dengan tipe data *Character*, umur dengan tipe data *numeric*, jenis kelamin dengan tipe data *Character*, diagnose penyakit dengan tipe data *Character*, dan jenis malaria dengan tipe data *Character*.

1.1.2. Seleksi Data Atribut

Seleksi data atribut merupakan proses penyeleksian data sesuai pada bab 4 yaitu beberapa pertanyaan seperti umur, alamat, dan diagnosa

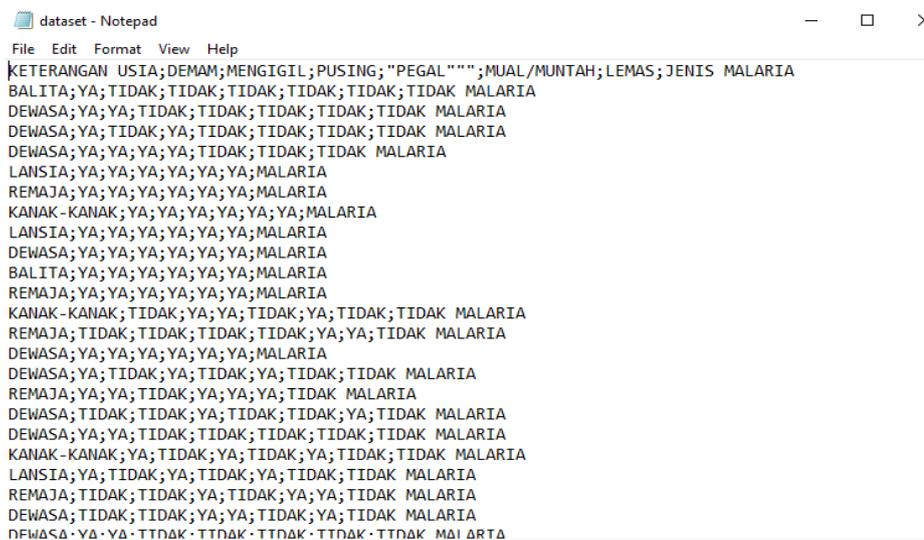
penyakit dan jenis malaria. Atribut selain yang disebutkan diatas akan dihilangkan dari data yang nantinya akan diolah. Masing-masing atribut tersebut memiliki item yang berbeda. Data seleksi data disajikan di tabel 5.2 :

Tabel 5.2 Seleksi Data Atribut

No	Nama Atribut	Action
1	Memiliki bpjs	Hapus atribut
2.	Memiliki Kartu Indonesia Sehat	Hapus atribut
3.	Jenis Kelamin	Hapus atribut

1.1.3. Konversi Hasil Pelabelan Data

Seluruh data yang diatas berbentuk *file excel*, yang kemudian akan disimpan kedalam bentuk ekstensi *file .csv (comma delimited)*. Setelah disimpan dalam bentuk *.csv (comma delimited) file* tersebut dibuka dengan *tools Weka*, untuk dikonversikan menjadi *file* ber-ekstensi *.arff*. Adapun bentuk data bereksentsi *.arff* tersebut apabila dibuka dengan aplikasi *notepad*, tampilan *arff* disajikan pada gambar 5.1:

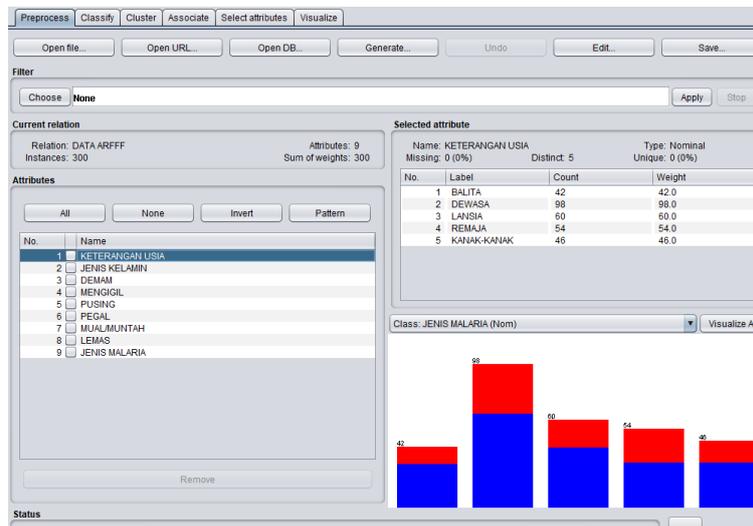


```
dataset - Notepad
File Edit Format View Help
KETERANGAN USIA;DEMAM;MENGIGIL;PUSING;"PEGAL""";MUAL/MUNTAH;LEMAS;JENIS MALARIA
BALITA;YA;TIDAK;TIDAK;TIDAK;TIDAK;TIDAK;TIDAK;TIDAK MALARIA
DEWASA;YA;YA;TIDAK;TIDAK;TIDAK;TIDAK;TIDAK;TIDAK MALARIA
DEWASA;YA;TIDAK;YA;TIDAK;TIDAK;TIDAK;TIDAK;TIDAK MALARIA
DEWASA;YA;YA;YA;YA;TIDAK;TIDAK;TIDAK;TIDAK MALARIA
LANSLA;YA;YA;YA;YA;YA;YA;MALARIA
REMAJA;YA;YA;YA;YA;YA;YA;MALARIA
KANAK-KANAK;YA;YA;YA;YA;YA;YA;MALARIA
LANSLA;YA;YA;YA;YA;YA;YA;MALARIA
DEWASA;YA;YA;YA;YA;YA;YA;MALARIA
BALITA;YA;YA;YA;YA;YA;YA;MALARIA
REMAJA;YA;YA;YA;YA;YA;YA;MALARIA
KANAK-KANAK;TIDAK;YA;YA;TIDAK;YA;TIDAK;TIDAK MALARIA
REMAJA;TIDAK;TIDAK;TIDAK;TIDAK;YA;TIDAK MALARIA
DEWASA;YA;YA;YA;YA;YA;YA;MALARIA
DEWASA;YA;TIDAK;YA;TIDAK;YA;TIDAK;TIDAK MALARIA
REMAJA;YA;YA;TIDAK;YA;YA;YA;TIDAK MALARIA
DEWASA;TIDAK;TIDAK;YA;TIDAK;TIDAK;YA;TIDAK MALARIA
DEWASA;YA;YA;TIDAK;TIDAK;TIDAK;TIDAK;TIDAK MALARIA
KANAK-KANAK;YA;TIDAK;YA;TIDAK;YA;TIDAK;TIDAK MALARIA
LANSLA;YA;TIDAK;YA;TIDAK;YA;TIDAK;TIDAK MALARIA
REMAJA;TIDAK;TIDAK;YA;TIDAK;YA;YA;TIDAK MALARIA
DEWASA;TIDAK;TIDAK;YA;YA;TIDAK;YA;TIDAK MALARIA
DEWASA;YA;YA;TIDAK;TIDAK;TIDAK;TIDAK;MALARIA
```

Gambar 5.1 Konversi Hasil Pelabelan Data .arff

1.2. HASIL VISUALISASI ATRIBUT DENGAN MENGGUNAKAN TOOLS WEKA

Berikut merupakan visualisasi menggunakan *tools* WEKA dari beberapa atribut Sseperti gambar 5.2:

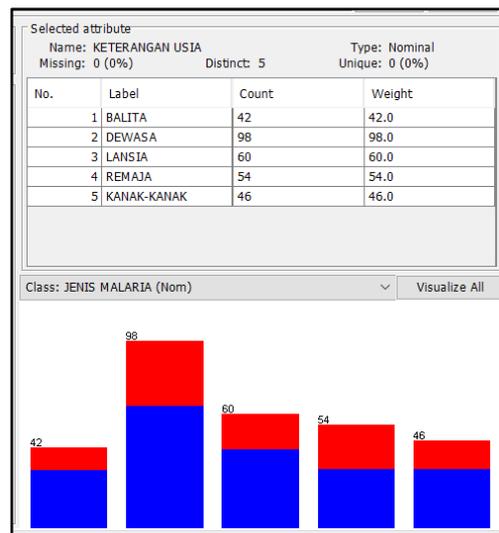


Gambar 5.2 Visualisai Atribut

Dari gambar 5.2 ditampilkan semua data yang dijadikan format arff dimasukkan ke *tools* untuk langkah visualisasi

1. Visualisasi Atribut Keterangan Usia

Visualisasi untuk melihat grafik umur ditampilkan pada gambar 5.3 yang menggunakan *tools* weka 3.9.

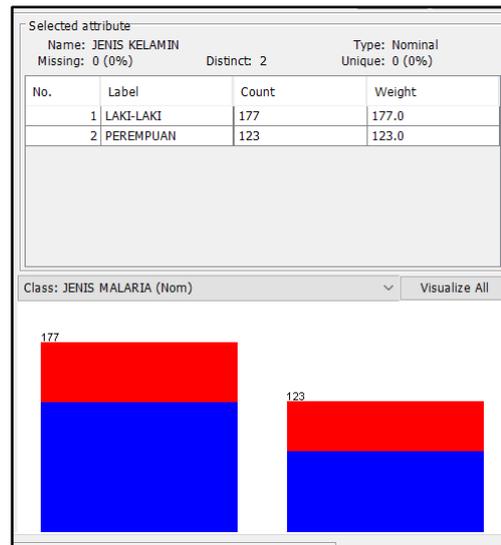


Gambar 5.3 Visualisasi Atribut Keterangan Usia

Gambar 5.3 adalah visualisasi dari atribut umur yang mana dari 300 *record* terdiri dari 5 *label* dan untuk nilai missing atau kesalahan adalah 0%, dengan terdiri dari 42 pasien yang masuk kedalam kategori balita, 98 pasien kategori dewasa, 60 pasien kategori lansia, 54 pasien kategori remaja dan 46 pasien kategori kanak-kanak, maka dapat disimpulkan bahwa dari atribut umur dengan *record* 300 rata-rata pasien masuk kedalam kategori dewasa.

2. Visualisasi Atribut Jenis Kelamin

Visualisasi untuk melihat grafik jenis kelamin ditampilkan pada gambar 5.4 yang menggunakan *tools* weka 3.9.

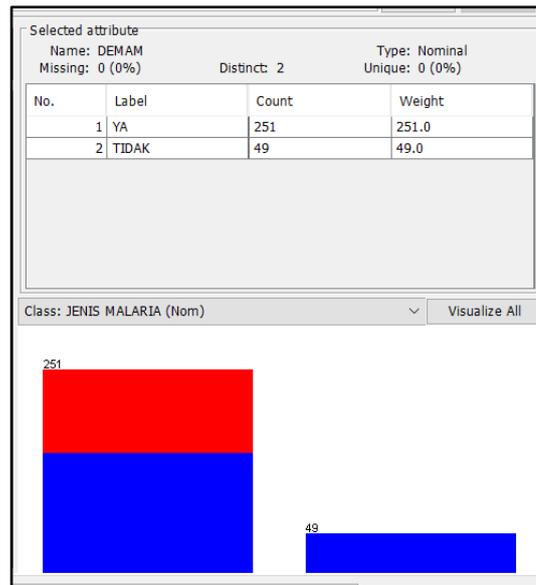


Gambar 5.4 Visualisasi Atribut Jenis Kelamin

Gambar 5.4 adalah visualisasi dari atribut jenis kelamin yang mana dari 300 *record* terdiri dari 2 *label* dan untuk nilai missing atau kesalahan adalah 0%, yang terdiri dari 177 pasien berjenis laki-laki dan 123 pasien berjenis perempuan. Dapat disimpulkan bahwa rata-rata jenis kelamin pasien adalah jenis kelamin laki-laki.

3. Visualisasi Atribut Demam

Visualisasi untuk melihat grafik demam ditampilkan pada gambar 5.5 yang menggunakan *tools* weka 3.9.

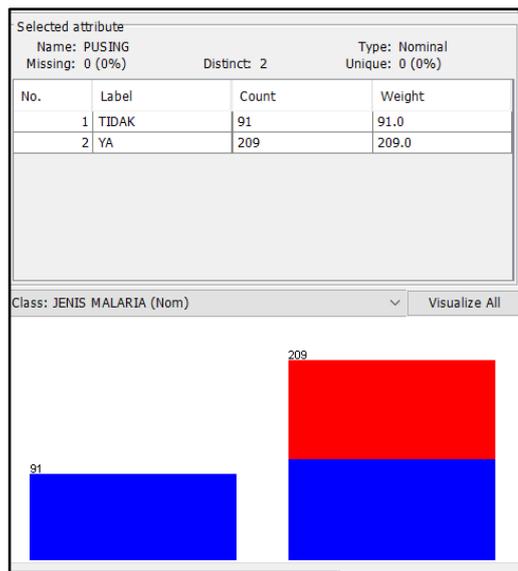


Gambar 5.5 Visualisasi Atribut Demam

Gambar 5.5 adalah visualisasi dari atribut demamyang mana dari 300 *record* terdiri dari 2 *label* dan untuk nilai missing atau kesalahan adalah 0%, yang terdiri dari 251 pasien yang mengalami demam dan 49 pasien tidak mengalami demam. Dapat disimpulkan bahwa pasien banyak mengalami demam dibandingkan dengan yang tidak demam.

4. Visualisasi Atribut Pusing

Visualisasi untuk melihat grafik pusing ditampilkan pada gambar 5.6 yang menggunakan *tools* weka 3.9.

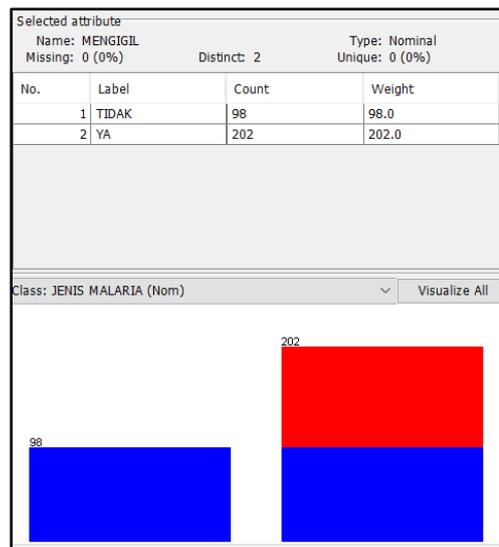


Gambar 5.6 Visualisasi Atribut Pusing

Gambar 5.6 adalah visualisasi dari atribut pusing yang mana dari 300 *record* terdiri dari 2 *label* dan untuk nilai missing atau kesalahan adalah 0%, yang terdiri dari 209 pasien yang mengalami pusing dan 91 pasien tidak mengalami pusing. Dapat disimpulkan bahwa pasien banyak mengalami pusing dibandingkan dengan yang tidak pusing.

5. Visualisasi Atribut Mengigil

Visualisasi untuk melihat grafik mengigil ditampilkan pada gambar 5.7 yang menggunakan *tools* weka 3.9



Gambar 5.7 Visualisasi Atribut Mengigil

Gambar 5.7 adalah visualisasi dari atribut mengigil yang mana dari 300 *record* terdiri dari 2 *label* dan untuk nilai missing atau kesalahan adalah 0%, yang terdiri dari 202 pasien yang mengalami mengigil dan 98 pasien tidak mengalami mengigil. Dapat disimpulkan bahwa pasien banyak mengalami mengigil dibandingkan dengan yang tidak mengigil.

6. Visualisasi Atribut Pegal-pegal

Visualisasi untuk melihat grafik pegal-pegal ditampilkan pada gambar 5.8 yang menggunakan *tools* weka 3.9

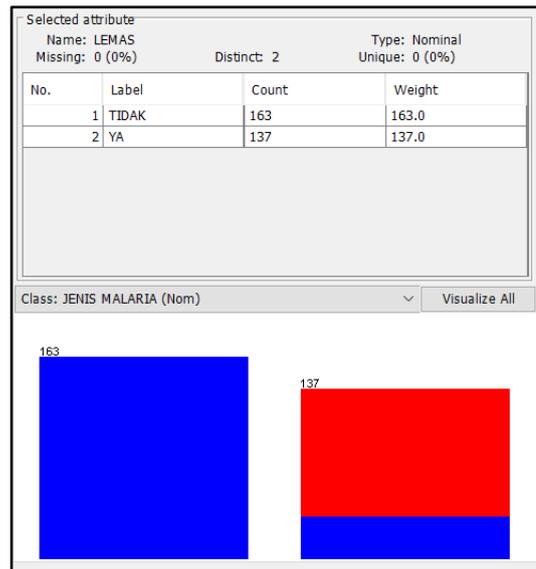


Gambar 5.8 Visualisasi Atribut Pegal

Gambar 5.8 adalah visualisasi dari atribut pegal yang mana dari 300 *record* terdiri dari 2 *label* dan untuk nilai missing atau kesalahan adalah 0%, yang terdiri dari 161 pasien yang mengalami pegal-pegal dan 138 pasien tidak mengalami pegal-pegal. Dapat disimpulkan bahwa pasien banyak mengalami pegal-pegal dibandingkan dengan yang tidak pegal-pegal.

7. Visualisasi Atribut Lemas

Visualisasi untuk melihat grafik lemas ditampilkan pada gambar 5.9 yang menggunakan *tools* weka 3.9

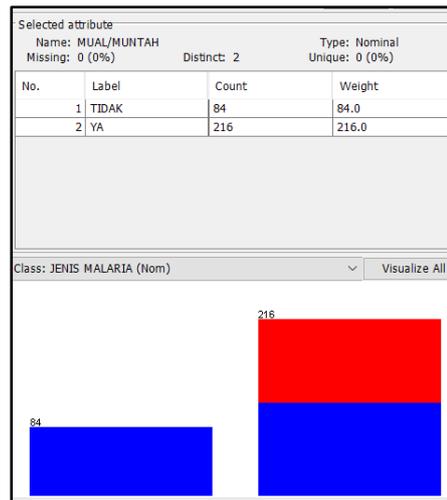


Gambar 5.9 Visualisasi Atribut Lemas

Gambar 5.9 adalah visualisasi dari atribut lemas yang mana dari 300 *record* terdiri dari 2 *label* dan untuk nilai missing atau kesalahan adalah 0%, yang terdiri dari 137 pasien yang mengalami lemas dan 163 pasien tidak mengalami lemas. Dapat disimpulkan bahwa pasien banyak mengalami tidak lemas dibandingkan dengan yang mengalami lemas.

8. Visualisasi Atribut Mual/Muntah

Visualisasi untuk melihat grafik lemas ditampilkan pada gambar 5.10 yang menggunakan *tools* weka 3.9



Gambar 5.10 Visualisasi Atribut Mual/Muntah

Gambar 5.10 adalah visualisasi dari atribut mual yang mana dari 300 *record* terdiri dari 2 *label* dan untuk nilai missing atau kesalahan adalah 0%, yang terdiri dari 216 pasien yang mengalami mual dan 84 pasien tidak mengalami mual. Dapat disimpulkan bahwa pasien banyak mengalami mual dibandingkan dengan yang mengalami mual.

9. Visualisasi Atribut Hasil Kondisi

Visualisasi untuk melihat grafik lemas ditampilkan pada gambar 5.11 yang menggunakan *tools* weka 3.9



Gambar 5.11 Visualisasi Atribut Hasil Kondisi

Gambar 5.11 adalah visualisasi dari atribut hasil kondisi yang mana dari 300 *record* terdiri dari 2 *label* dan untuk nilai missing atau kesalahan adalah 0%, yang terdiri dari 103 pasien malaria dan 197 pasien tidak malaria. Dapat disimpulkan bahwa pasien banyak mengalami malaria dibandingkan dengan yang tidak malatia.

1.2.1. Hasil Klasifikasi *Decision Tree* Dengan *Tools Weka*

Klasifikasi menggunakan *Decision Tree* dilakukan dengan menggunakan *tools Weka*.

1. Hasil Klasifikasi *Use Data Training*

Test Use Data Traaining merupakan data menggunakan *data training* itu sendiri, dapat dilihat pada tabel 5.3:

Tabel 5.3 Summary Use Data Training Decision Tree Dengan Tools Weka

<i>Data Summary</i>	<i>Nilai</i>	<i>Persentase</i>
<i>Correctly Classified Instances</i>	299	99.6667%
<i>Incorrectly Classified Instances</i>	1	0.3333%
<i>Kappa statistic</i>	0.9926	-
<i>Mean absolute error</i>	0.0066	-
<i>Root mean squared error</i>	0.0575	-
<i>Relative absolute error</i>	1.4632%	-
<i>Root relative squared error</i>	12.1007%	-
<i>Total Number of Instances</i>	300	-

Berdasarkan hasil dari tabel 5.3 adalah *Instance* yang Diklasifikasikan dengan Benar 299 adalah 99.6667%, *Instance* yang Diklasifikasikan dengan Salah 1 adalah 0.3333 %, *Statistik Kappa* 0.9926, dengan kesalahan absolut 0,0066, *Root mean* kuadrat kesalahan 0,0575, Kesalahan absolut relative 1.4632 %, Kesalahan kuadrat relatif akar 12.1007%Jumlah Total Instance 300 data. Dapat disimpulkan bahwa perhitungan dengan *use data training* menggunakan *tools* Weka prediksi benar lebih tinggi di bandingkan dengan prediksi salah. Untuk tabel *confusion matrix* dapat dilihat pada tabel 5.4, *confusion matrix* adalah tabel yang memberikan informasi perbandingan hasil klasifikasi yang dilakukan oleh system dengan hasil klasifikasi yang sebenarnya.

Tabel 5.4 Confusion Matrix Use Training Decision Tree Dengan Tools Weka

<i>Classified as</i>	Malaria	Tidak Malaria
A	196	1
B	0	103

Berdasarkan tabel 5.4 nilai akhir dari *Decision Tree* adalah *classified as* A untuk malaria 196, dan tidak malaria adalah 1, selanjutnya *classified as* B untuk malaria 0, dan tidak malaria adalah 103. Dapat di simpulkan bahwa dari 196 data

yang tepat dari kelas tidak malaria memiliki 1 data *miss calculate* dan 0 data yang tepat dari kelas malaria memiliki 103 data *miss calculate*.

2. *Cross Validation Folds 10* pada *Decision Tree*

Test 10 *Cross Validation* melakukan pengtesan data dimana data *training* dibagi menjadi k buah *subset* ((sub himpunan). Dimana k adalah nilai dari *fold*, pada pengtesan ini nilai *fold* adalah 10. Selanjutnya untuk tiap dari *subset*, akan dijadikan data tes dari hasil klasifikasi yang dihasilkan dari k-1 *subset* lainnya. Jadi, akan ada 10 kali tes. Dimana setiap data akan menjadi data tes sebanyak 1 kali dan data *training* sebanyak k-1 kali. Hasil dari *Cros Validation Flods10* dapat dilihat pada tabel 5.5:

Tabel 5.5 Summary Cros Validation 10 Klasifikasi Decision Tree Dengan Tools Weka

<i>Data Summary</i>	Nilai	Persentase (%)
<i>Correctly Classified Instances</i>	299	99.6667%
<i>Incorrectly Classified Instances</i>	1	0.3333%
<i>Kappa statistic</i>	0.9926	-
<i>Mean absolute error</i>	0.0063	-
<i>Root mean squared error</i>	0.059	-
<i>Relative absolute error</i>	1.3861%	-
<i>Root relative squared error</i>	12.4269%	-
<i>Total Number of Instances</i>	300	-

Berdasarkan hasil dari tabel 5.5 adalah *Instance* yang Diklasifikasikan dengan Benar 299 adalah 99.6667%, *Instance* yang Diklasifikasikan dengan Salah 1 adalah 0.3333 %, *Statistik Kappa* 0.9926, Berarti kesalahan absolut 0.0063, *Root mean* kuadrat kesalahan 0,059, Kesalahan absolut relative 1.3861 %, Kesalahan kuadrat relatif akar 12.4269%Jumlah Total Instance 300 data. . Untuk tabel *confusion matrix* dapat dilihat pada tabel 5.6, *confusion matrix* adalah tabel yang

memberikan informasi perbandingan hasil klasifikasi yang dilakukan oleh system dengan hasil klasifikasi yang sebenarnya.

Tabel 5.6 Confusion Matrix Cros Validation 10 Decision Tree Dengan Tools Weka

<i>Classified as</i>	Malaria	Tidak Malaria
A	0	196
B	103	1

Berdasarkan tabel 5.6 nilai akhir dari *Decision Tree* adalah *classified as A* untuk malaria 0, dan tidak malaria adalah 196, selanjutnya *classified as B* untuk malaria 103, dan tidak malaria adalah 1. Dapat di simpulkan bahwa dari 0 data yang tepat dari kelas tidak malaria memiliki 196 data *miss calculate* dan 103 data yang tepat dari kelas malaria memiliki 1 data *miss calculate*.

3. *Cros Validation 5* pada *Decision Tree*

Test 5 *Cross Validation* melakukan pengetesan data dimana data *training* dibagi menjadi k buah *subset* ((sub himpunan). Dimana k adalah nilai dari *fold*, pada pengetesan ini nilai *fold* adalah 5. Selanjutnya untuk tiap dari *subset*, akan dijadikan data tes dari hasil klasifikasi yang dihasilkan dari k-1 *subset* lainnya. Jadi, akan ada 5 kali tes. Dimana setiap data akan menjadi data tes sebanyak 1 kali dan data *training* sebanyak k-1 kali Hasil dari *Cros Validation folds 5*, disajikan pada tabel 5.6:

Tabel 5.7 Sumary Cros Validation 5 Decision Tree Dengan Tools Weka

<i>Data Sumary</i>	Nilai	Persentase (%)
<i>Correctly Classified Instances</i>	298	99.3333%
<i>Incorrectly Classified Instances</i>	2	0.6667%
<i>Kappa statistic</i>	0.9853	-
<i>Mean absolute error</i>	0.0097	-
<i>Root mean squared error</i>	0.0822	-
<i>Relative absolute error</i>	2.1549%	-

<i>Root relative squared error</i>	17.3026%	-
<i>Total Number of Instances</i>	300	-

Berdasarkan hasil dari tabel 5.7 adalah *Instance* yang Diklasifikasikan dengan Benar 298 adalah 99.3333%, *Instance* yang Diklasifikasikan dengan Salah 2 adalah 0.6667 %, *Statistik Kappa* 0.9853, Berarti kesalahan absolut 0,0097, *Root mean* kuadrat kesalahan 0,0822, Kesalahan absolut relative 2.1549 %, Kesalahan kuadrat relatif akar 17.3026%Jumlah Total *Instance* 300 data. Untuk tabel *confusion matrix* dapat dilihat pada tabel 5.8, *confusion matrix* adalah tabel yang memberikan informasi perbandingan hasil klasifikasi yang dilakukan oleh system dengan hasil klasifikasi yang sebenarnya.

Tabel 5.8 Confusion Matrix Cros Validation 5 Decision Tree Dengan Tools Weka

<i>Classified as</i>	Malaria	Tidak Malaria
A	0	195
B	103	2

Berdasarkan tabel 5.8 nilai akhir dari *Decision Tree* adalah *classified as A* untuk malaria 0, dan tidak malaria adalah 195 selanjutnya *classified as B* untuk malaria 195, dan tidak malaria adalah 2. Dapat di simpulkan bahwa dari 0 data yang tepat dari kelas tidak malaria memiliki 195 data *miss calculate* dan 103 data yang tepat dari kelas malaria memiliki 2 data *miss calculate*.

4. *Percentage split 66%* pada *Decision Tree*

Test 60% *percentage split* merupakan pengetesan data dimana data *training* dibagi menjadi k buah *subset* (sub himpunan). Dimana k adalah nilai dari *percentage*. Pada pengetesan ini nilai *percentage* 60%, selanjutnya untuk tiap dari

subset, akan dijadikan data tes dari hasil klasifikasi yang dihasilkan dari k-1 *subset* lainnya. Hasil dari *Percentage split 66%*, disajikan pada tabel 5.9:

Tabel 5.9 Percentage split 66% Decision Tree Dengan Tools Weka

<i>Data Summary</i>	Nilai	<i>Persentase (%)</i>
<i>Correctly Classified Instances</i>	102	100%
<i>Incorrectly Classified Instances</i>	0	0%
<i>Kappa statistic</i>	1	-
<i>Mean absolute error</i>	0.0054	-
<i>Root mean squared error</i>	0.009	-
<i>Relative absolute error</i>	1.1907%	-
<i>Root relative squared error</i>	1.8666%	-
<i>Total Number of Instances</i>	102	-

Berdasarkan hasil dari tabel 5.9 adalah Instance yang Diklasifikasikan dengan Benar 102 adalah 100%, Instance yang Diklasifikasikan dengan Salah 0 adalah 0%, Statistik Kappa 1, Berarti kesalahan absolut 0,0054, Root mean kuadrat kesalahan 0,009, Kesalahan absolut relatif 1.1907 %, Kesalahan kuadrat relatif akar 1.8666 %, Jumlah Total Instance 102 data. Untuk tabel *confusion matrix* dapat dilihat pada tabel 5.9, *confusion matrix* adalah tabel yang memberikan informasi perbandingan hasil klasifikasi yang dilakukan oleh system dengan hasil klasifikasi yang sebenarnya

Tabel 5.9 Confusion Matrix Percentage split 66% Decision Tree Dengan Tools Weka

<i>Classified as</i>	Malaria	Tidak Malaria
A	65	0
B	0	37

Berdasarkan tabel 5.9 nilai akhir dari *Decision Tree* adalah *classified as A* untuk malaria 65, dan tidak malaria adalah 0, selanjutnya *classified as B* untuk malaria 0, dan tidak malaria adalah 37. Dapat di simpulkan bahwa dari 65 data

yang tepat dari kelas tidak malaria memiliki 0 data *miss calculate* dan 0 data yang tepat dari kelas malaria memiliki 37 data *miss calculate*.

1.2.2. Perbandingan Hasil Klasifikasi Algoritma C4.5

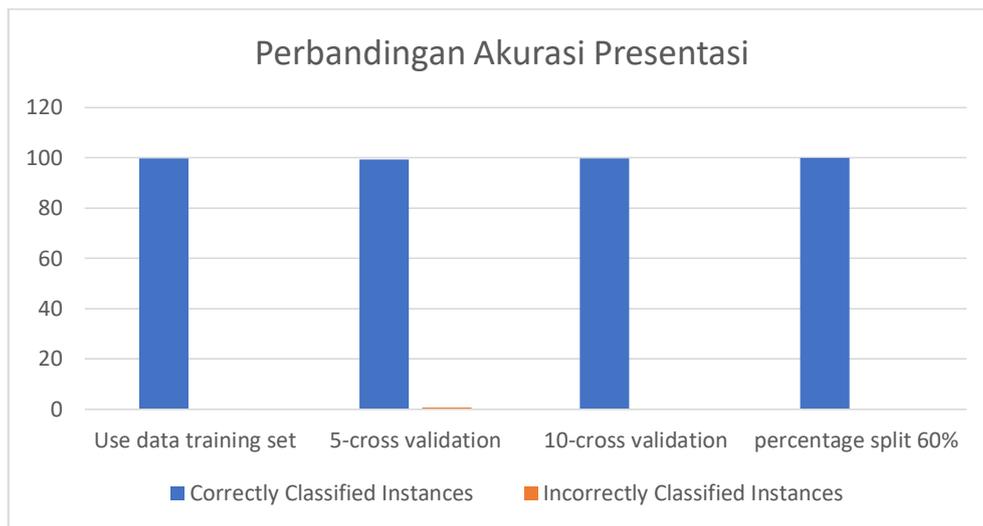
Setelah dilakukan analisis klasifikasi Algoritma C4.5 pada *tools* Weka dengan 4 *test* yaitu *use data training*, *5-cross validation*, *10-cross validation* dan *percentage split 60%*. Maka diperoleh akurasi tertinggi yaitu menggunakan *percentage split 60%* dan *5-cross validation* dengan akurasi presentasi sebesar 100% untuk *Correctly Classified Instances* dan 0.667% untuk *Incorrectly Classified Instances*. Hasil perbandingan di tampilkan pada tabel 5.10.

Tabel 5.10 Hasil Perbandingan Akurasi Presentasi

	Akurasi	Jumlah Data	Presentasi (%)
<i>Use data training set</i>	<i>Correctly Classified Instances</i>	299	99.667%
	<i>Incorrectly Classified Instances</i>	1	0.3333%
<i>5-cross validation</i>	<i>Correctly Classified Instances</i>	298	99.3333%
	<i>Incorrectly Classified Instances</i>	2	0.667%
<i>10-cross validation</i>	<i>Correctly Classified Instances</i>	299	99.667%
	<i>Incorrectly Classified Instances</i>	1	0.3333%
<i>percentage split 60%</i>	<i>Correctly Classified Instances</i>	120	100%
	<i>Incorrectly Classified Instances</i>	0	0%

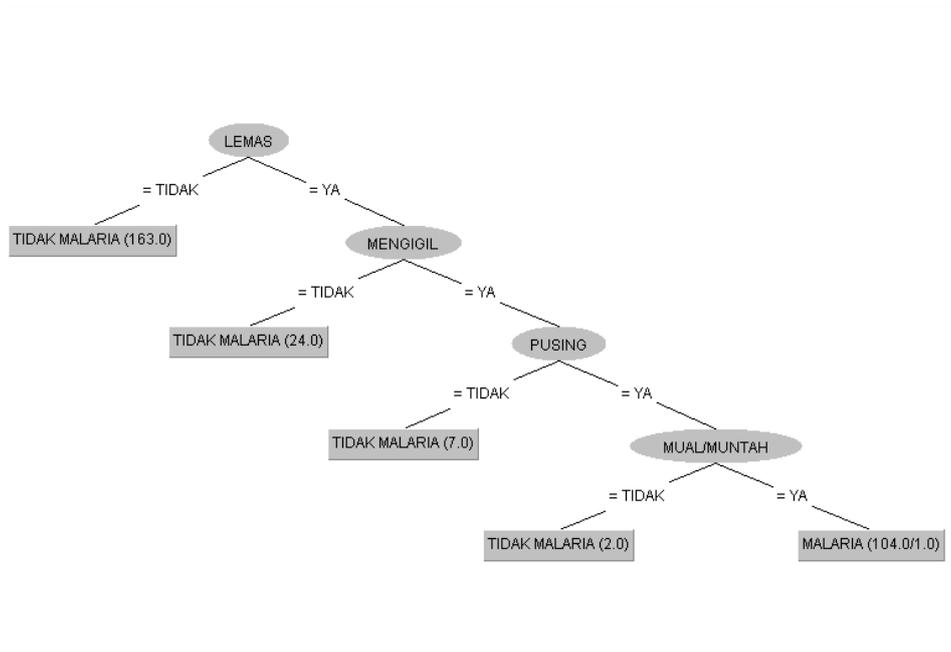
Dari tabel 5.10 perbandingan 4 *test* menghasilkan presentasi akurasi yang untuk *use data training* dengan akurasi benar 99.667% dan presentasi tidak tepat

0.3333%, 5-cross validation dengan akurasi benar 99.3333% dan presentasi tidak tepat 0.667%, 10-cross validation dengan akurasi benar 99.667% dan presentasi tidak tepat 0.3333% dan percentage split 60% dengan akurasi benar 100% dan presentasi tidak tepat 0%. Dari presentasi pada tabel 5.10 digambarkan grafik yang di tampilkan pada gambar 5.12



Gambar 5.12 presentasi akurasi klasifikasi C4.5

Dari gambar 5.12 didapat kesimpulan bahwa prediksi benar lebih banyak dibandingkan prediksi salah, maka dari itu dari hasil klasifikasi dengan *tools* weka di peroleh sebuah pohon keputusan yang ditampilkan pada gambar 5.13.



Gambar 5.13 Pohon Keputusan Menggunakan Tools Weka

Dari gambar 5.13 yang mana bagan itu menunjukkan bagan sebuah alur keputusan dengan *rols*

1. Jika **lemas** = **tidak** maka keputusan **tidak malaria**
2. Jika **lemas** = **ya** dan **mengigil** = **tidak** dan **pusing** = **tidak** maka keputusan **tidak malaria**
3. Jika **lemas** = **ya** dan **mengigil** = **tidak** maka keputusan **tidak malaria**
4. Jika **lemas** = **ya** dan **mengigil** = **ya** dan **pusing** = **ya** dan **mual** = **ya** maka keputusan **malaria**
5. Jika **lemas** = **ya** dan **mengigil** = **ya** dan **pusing** = **ya** dan **mual** = **tidak** maka keputusan **tidak malaria**