

## BAB V

### HASIL ANALISIS DAN VISUALISASI

#### 5.1 HASIL ANALISIS DAN VISUALISASI DATA MENGGUNAKAN *TOOLS WEKA*

Pada tahapan ini adalah analisis dan visualisasi data. Berikut adalah hasil analisis dan visualisasi data menggunakan tools weka :

##### 5.1.1 *Data Experimen*

Berikut adalah tampilan dari data experimen sebelum diseleksi

**Tabel 5.1 *Data Experimen***

No	Umur	Jenis Kelamin	Alamat	Status Kunjungan	Class Diagnosa Hipertensi
1	61	Perempuan	Sei.Landai	Baru	Negatif
2	29	Perempuan	Punggur	Baru	Negatif
3	26	Perempuan	Punggur	Baru	Negatif
4	50	Perempuan	Punggur	Lama	Negatif
5	53	Perempuan	Bor 3	Lama	Negatif

##### 5.1.2 *Seleksi Data Experimen*

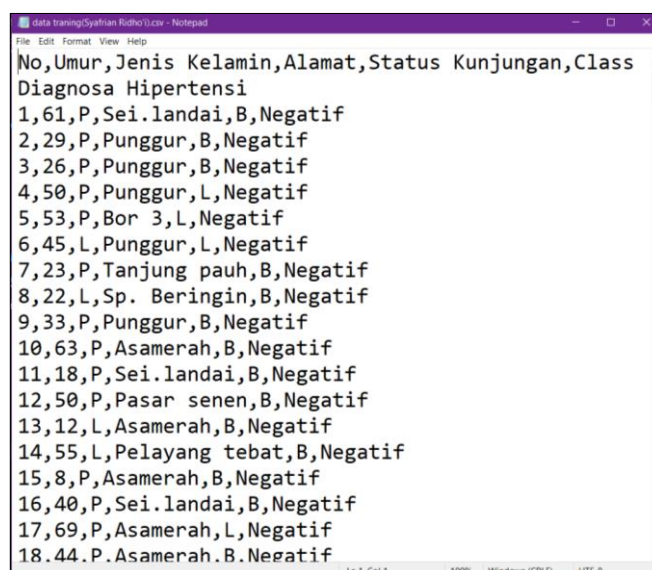
Seleksi *data experimen* dilakukan penulis untuk data tidak ada *noise* dan data pasien pada tahun 2019-2020 yang penulis gunakan untuk penelitian penulis. Berikut adalah hasil seleksi data experimen yang diambil dari atribut umur, jenis kelamin, alamat, status kunjungan dan *class diagnose hipertensi*.

Tabel 5.2 Seleksi *Data Experimen*

No	Umur	Jenis Kelamin	Alamat	Status Kunjungan	Class Diagnosa Hipertensi
1	4	Perempuan	Punggur	Baru	Negatif
2	7	Perempuan	Pelayangan tebat	Baru	Negatif
3	7	Laki-laki	Asamerah	Baru	Negatif
4	7	Laki-laki	Lubuk Penyengat	Lama	Negatif
5	7	Perempuan	Bor 3	Lama	Negatif

### 5.1.3 Konversi Hasil Pelabelan Data

Berikut adalah hasil dari konversi data dari file *excel* diubah menjadi file *.csv* (*comma delimited*), kemudian diolah dengan *tools Weka*, untuk dikonversikan menjadi *file* ber-ekstensi *.arff*. Adapun bentuk data berekstensi *.arff* tersebut apabila dibuka dengan aplikasi *notepad* adalah sebagai berikut:



```

data training(Syahrin Ridho)\.csv - Notepad
File Edit Format View Help
No,Umur,Jenis Kelamin,Alamat,Status Kunjungan,Class
Diagnosa Hipertensi
1,61,P,Sei.landai,B,Negatif
2,29,P,Punggur,B,Negatif
3,26,P,Punggur,B,Negatif
4,50,P,Punggur,L,Negatif
5,53,P,Bor 3,L,Negatif
6,45,L,Punggur,L,Negatif
7,23,P,Tanjung pauh,B,Negatif
8,22,L,Sp. Beringin,B,Negatif
9,33,P,Punggur,B,Negatif
10,63,P,Asamerah,B,Negatif
11,18,P,Sei.landai,B,Negatif
12,50,P,Pasar senen,B,Negatif
13,12,L,Asamerah,B,Negatif
14,55,L,Pelayang tebat,B,Negatif
15,8,P,Asamerah,B,Negatif
16,40,P,Sei.landai,B,Negatif
17,69,P,Asamerah,L,Negatif
18.44.P.Asamerah.B.Negatif
Ln 1, Col 1      100%  Windows (CRLF)  UTF-8

```

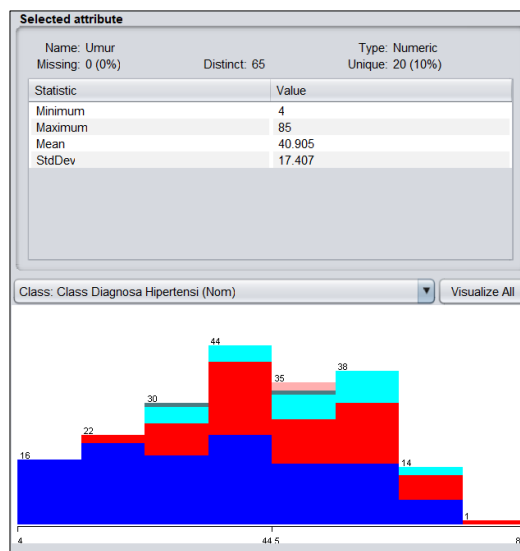
Gambar 5.1 Konversi Hasil Pelabelan Data *.arff*

## 5.2 BENTUK VISUALISASI DENGAN MENGGUNAKAN *TOOLS WEKA*

### 5.2.1 Hasil Visualisasi Atribut Dengan Menggunakan *Tools Weka*

Berikut merupakan bentuk visualisasi menggunakan *tools Weka* dari beberapa atribut yaitu sebagai berikut :

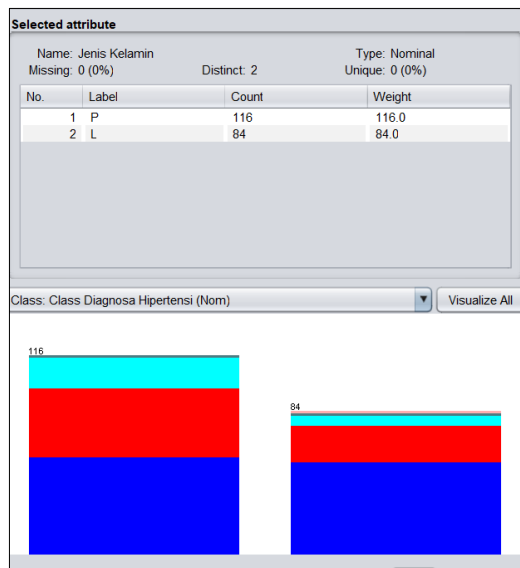
#### 1. Visualisasi Atribut Umur



**Gambar 5.2 Visualisasi Atribut Umur**

Berdasarkan gambar 5.2 adalah visualisasi dari atribut umur memiliki nilai statistik umur minimum adalah 4, nilai maximum adalah 85 sehingga memiliki nilai mean adalah 40,905 dan untuk nilai missing atau kesalahan adalah 0%.

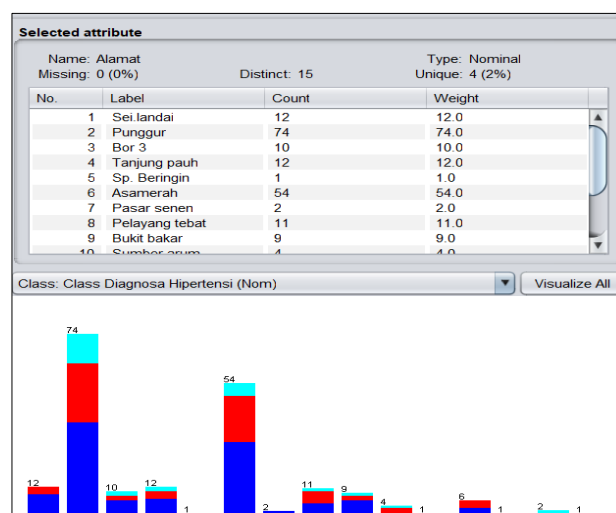
## 2. Visualisasi Atribut Jenis Kelamin



**Gambar 5.3 Visualisasi Atribut Jenis Kelamin**

Berdasarkan gambar 5.3 adalah visualisasi dari atribut jenis kelamin P berjumlah 116, jenis kelamin L berjumlah 84, dan untuk nilai missing atau kesalahan adalah 0%.

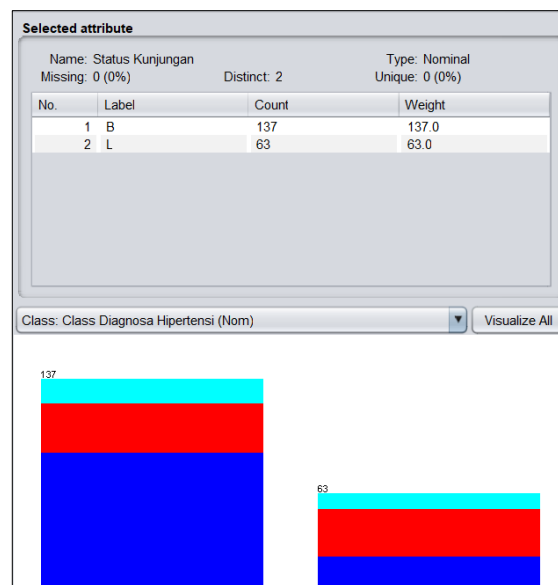
## 3. Visualisasi Atribut Alamat



**Gambar 5.4 Visualisasi Atribut Alamat**

Pada gambar 5.4 visualisasi atribut 15 alamat yang berbeda tiap pasien yaitu sei landau berjumlah 12, punggung berjumlah 74, bor 3 berjumlah 10, tanjung pauh berjumlah 12, sp beringin berjumlah 1, asamerah berjumlah 54, pasar senen berjumlah 2. Pelayang tebat berjumlah 11, bukit bakar berjumlah 9, sumber arum berjumlah 4, kelapa kembar berjumlah 1, kilis berjumlah 6, bukit rinting 1, lubuk penyengat berjumlah 2, madrasah berjumlah 1 dna memiliki nilai missing atau kesalahan pada alamat adalah 0%.

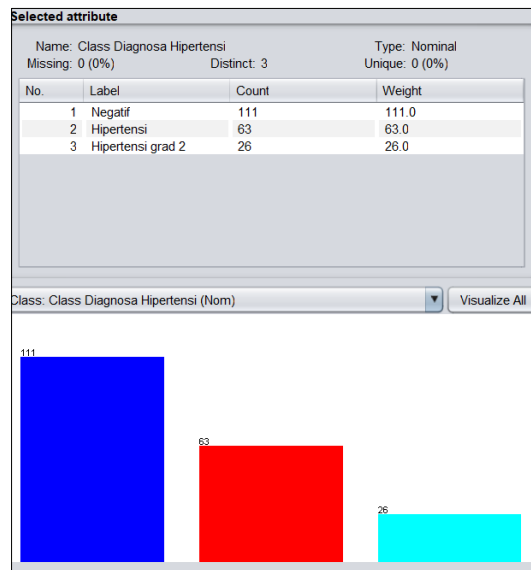
#### 4. Visualisasi Atribut Status Kunjungan



**Gambar 5.5 Visualisasi Atribut Status Kunjungan**

Pada gambar 5.5 visualisasi atribut memiliki 2 label yaitu *b* berjumlah 137 dan *l* berjumlah 63, dengan nilai kesalahan atau *missing* 0 %.

## 5. Visualisasi Atribut *Class Diagnosa Hipertensi*



**Gambar 5.6** Visualisasi Atribut *Class Diagnosa Hipertensi*

Pada gambar 5.6 visualisasi atribut memiliki 3 *label* yaitu negative berjumlah 111 , hipertensi 63, hipertensi grad 2 berjumlah 26, dan yes, dengan nilai kesalahan atau *missing* 0 %.

### 5.2.2 Hasil Klasifikasi *Native Bayes* Dengan *Tools Weka*

Klasifikasi menggunakan *native bayes* dilakukan dengan menggunakan *tools Weka*.

1. *Cros Validation Folds* 10 pada Klasifikasi *Native Bayes*

Berikut adalah hasil dari *Cros Validation Flods*10.

**Tabel 5.3 Summary Cros Validation 10 Klasifikasi Native Bayes Dengan Tools****Weka**

<i>Data Summary</i>	<i>Nilai</i>	<i>Persentase</i>
<i>Correctly Classified Instances</i>	110	59%
<i>Incorrectly Classified Instances</i>	82	41%
<i>Kappa statistic</i>	0.2298	-
<i>Mean absolute error</i>	0.3125	-
<i>Root mean squared error</i>	0.4072	-
<i>Relative absolute error</i>	81.1746 %	-
<i>Root relative squared error</i>	92.9065 %	-
<i>Total Number of Instances</i>	200	-

Berdasarkan hasil dari tabel 5.3 adalah *Instance* yang Diklasifikasikan dengan Benar 118 adalah 59%, *Instance* yang Diklasifikasikan dengan Salah 82 adalah 41 %, *Statistik Kappa* 0.2298, Berarti kesalahan absolut 0,3125, *Root mean* kuadrat kesalahan 0,4072, Kesalahan absolut relatif 81,1746 %, Kesalahan kuadrat relatif akar 92,9065 %Jumlah Total *Instance* 200 data.

**Tabel 5.4 Detailed Accuracy By Class Cros Validation 10 Klasifikasi Native****Bayes Dengan Tools Weka**

<i>TP Rate</i>	<i>FP Rate</i>	<i>Precision</i>	<i>Recall</i>	<i>F-Measure</i>	<i>MCC</i>	<i>ROC Area</i>	<i>PRC Area</i>	<i>Class</i>
0.766	0.483	0.664	0.766	0.711	0.293	0.780	0.843	Neaktif
0.524	0.277	0.465	0.524	0.493	0.239	0.719	0.513	Hipertensi
0.000	0.006	0.000	0.000	0.000	- 0.027	0.627	0.164	Hipertensi Grad 2
0.590	0.356	0.515	0.590	0.550	0.234	0.741	0.651	<i>Weighted Avg.</i>

Berdasarkan tabel 5.4 rata-rata (*weighted avg.*) pada *detail accuracy by class cros validation folds 10* adalah nilai *TP Rate* adalah 0.590, nilai *FP Rate*

adalah 0.356, nilai *Precision* adalah 0.515, nilai *recall* adalah 0.590, nilai *f-measure* adalah 0.550, nilai *mcc* adalah 0.234, nilai *roc area* adalah 0.741 dan nilai *prc area* adalah 0.651.

**Tabel 5.5 Confusion Matrix Cros Validation 10 Klasifikasi Native Bayes  
Dengan Tools Weka**

<i>Classified as</i>	Negatif	Hipetensi	Hipertensi Grad 2
A	85	25	1
B	30	33	0
C	13	13	0

Berdasarkan tabel 5.5 nilai akhir dari native bayes adalah *classified as A* untuk negatif 88, hipertensi 25 dan hipertensi grad 2 adalah 1, selanjutnya *classified as B* untuk negatif 30, hipertensi 33, dan hipertensi grad 2 adalah 0, kemudian *classified as C* untuk negatif 13, hipertensi 13 dan hipertensi grad 2 adalah 0.

## 2. *Cros Validation 5* pada Klasifikasi Native Bayes

Berikut adalah hasil dari *Cros Validation folds 5*.

**Tabel 5.6 Sumary Cros Validation 5 Klasifikasi Native Bayes Dengan Tools  
Weka**

<i>Data Summary</i>	Nilai	<i>Persentase</i>
<i>Correctly Classified Instances</i>	117	58.5%
<i>Incorrectly Classified Instances</i>	83	41.5%
<i>Kappa statistic</i>	0.2178	-
<i>Mean absolute error</i>	0.3136	-
<i>Root mean squared error</i>	0.4094	-
<i>Relative absolute error</i>	81.4395%	-
<i>Root relative squared error</i>	93.4253%	-
<i>Total Number of Instances</i>	200	-



Berdasarkan hasil dari tabel 5.6 adalah Instance yang Diklasifikasikan dengan Benar 117 adalah 58.5%, Instance yang Diklasifikasikan dengan Salah 83 adalah 41.5 %, Statistik Kappa 0.2178, Berarti kesalahan absolut 0,3136, Root mean kuadrat kesalahan 0,4094, Kesalahan absolut relatif 81,4395 %, Kesalahan kuadrat relatif akar 93,4253 %Jumlah Total Instance 200 data.

**Tabel 5.7 Detailed Accuracy By Class Cros Validation 5 Klasifikasi Native  
Bayes Dengan Tools Weka**

<i>TP Rate</i>	<i>FP Rate</i>	<i>Precision</i>	<i>Recall</i>	<i>F-Measure</i>	<i>MCC</i>	<i>ROC Area</i>	<i>PRC Area</i>	<i>Class</i>
0.784	0.494	0.664	0.784	0.719	0.303	0.779	0.844	Neaktif
0.476	0.263	0.455	0.476	0.465	0.211	0.702	0.506	Hipertensi
0.000	0.017	0.000	0.000	0.000	- 0.048	0.633	0.170	Hipertensi Grad 2
0.585	0.359	0.512	0.585	0.546	0.228	0.736	0.650	<i>Weighted Avg.</i>

Berdasarkan tabel 5.7 rata-rata (*weighted avg.*) pada *detail accuracy by class cros validation folds 5* adalah nilai *TP Rate* adalah 0.585, nilai *FP Rate* adalah 0.359, nilai *Precision* adalah 0.512, nilai *recall* adalah 0.585, nilai *f-measure* adalah 0.546, nilai *mcc* adalah 0.228, nilai *roc area* adalah 0.736 dan nilai *prc area* adalah 0.650.

**Tabel 5.8 Confusion Matrix Cros Validation 5 Klasifikasi Native Bayes Dengan Tools Weka**

<i>Classified as</i>	Negatif	Hipetensi	Hipertensi Grad 2
A	87	23	1
B	31	30	2
C	13	13	0

Berdasarkan tabel 5.8 nilai akhir dari *Native Bayes* adalah classified as A untuk negatif 87, hipertensi 23 dan hipertensi grad 2 adalah 1, selanjutnya classified as B untuk negatif 31, hipertensi 30, dan hipertensi grad 2 adalah 2, kemudian classified as C untuk negatif 13, hipertensi 13 dan hipertensi grad 2 adalah 0.

3. *Use Training* pada Klasifikasi *Native Bayes*

Berikut adalah hasil dari *Use Training*:

**Tabel 5.9 Summary Use Training Klasifikasi Native Bayes Dengan Tools Weka**

<i>Data Summary</i>	Nilai	<i>Persentase</i>
<i>Correctly Classified Instances</i>	130	65%
<i>Incorrectly Classified Instances</i>	70	35%
<i>Kappa statistic</i>	0.3395	-
<i>Mean absolute error</i>	0.2907	-
<i>Root mean squared error</i>	0.3814	-
<i>Relative absolute error</i>	75.5359	-
<i>Root relative squared error</i>	87.0424	-
<i>Total Number of Instances</i>	200	-

Berdasarkan hasil dari tabel 5.9 adalah Instance yang Diklasifikasikan dengan Benar 130 adalah 65%, Instance yang Diklasifikasikan dengan Salah 70 adalah 35

%, Statistik Kappa 0.3395, Berarti kesalahan absolut 0,2907, Root mean kuadrat kesalahan 0,3814, Kesalahan absolut relatif 75,5359 %, Kesalahan kuadrat relatif akar 87,0424 %, Jumlah Total Instance 200 data.

**Tabel 5.10 Detailed Accuracy By Class Use Training Klasifikasi Native Bayes Dengan Tools Weka**

<i>TP Rate</i>	<i>FP Rate</i>	<i>Precision</i>	<i>Recall</i>	<i>F-Measure</i>	<i>MCC</i>	<i>ROC Area</i>	<i>PRC Area</i>	<i>Class</i>
0.829	0.427	0.708	0.829	0.763	0.419	0.830	0.879	Neaktif
0.587	0.234	0.536	0.587	0.561	0.346	0.782	0.627	Hipertensi
0.038	0.000	1.000	0.038	0.074	0.183	0.801	0.386	Hipertensi Grad 2
0.650	0.311	0.692	0.650	0.610	0.365	0.811	0.736	<i>Weighted Avg.</i>

Berdasarkan tabel 5.10 rata-rata (*weighted avg.*) pada *detail accuracy by class Use Training* adalah nilai *TP Rate* adalah 0.650, nilai *FP Rate* adalah 0.311, nilai *Precision* adalah 0.692, nilai *recall* adalah 0.650, nilai *f-measure* adalah 0.610, nilai *mcc* adalah 0.365, nilai *roc area* adalah 0.811 dan nilai *prc area* adalah 0.736.

**Tabel 5.11 Confusion Matrix Use Training Klasifikasi Native Bayes Dengan Tools Weka**

<i>Classified as</i>	Negatif	Hipetensi	Hipertensi Grad 2
A	92	19	0
B	26	37	0
C	12	13	1

Berdasarkan tabel 5.11 nilai akhir dari native bayes adalah classified as A untuk negatif 92, hipertensi 19 dan hipertensi grad 2 adalah 0, selanjutnya classified

as B untuk negatif 26, hipertensi 37, dan hipertensi grad 2 adalah 0, kemudian classified as C untuk negatif 12, hipertensi 13 dan hipertensi grad 2 adalah 1.

4. *Percentage split 66%* pada Klasifikasi *Native Bayes*

Berikut adalah hasil dari *Percentage split 66%*.

**Tabel 5.12 *Percentage split 66%* Klasifikasi *Native Bayes* Dengan *Tools Weka***

<i>Data Summary</i>	Nilai	<i>Persentase</i>
<i>Correctly Classified Instances</i>	46	67.6471%
<i>Incorrectly Classified Instances</i>	22	32.3529%
<i>Kappa statistic</i>	0.3725	-
<i>Mean absolute error</i>	0.2993	-
<i>Root mean squared error</i>	0.3953	-
<i>Relative absolute error</i>	78.5524 %	-
<i>Root relative squared error</i>	92.6317	-
<i>Total Number of Instances</i>	68	-

Berdasarkan hasil dari tabel 5.12 adalah Instance yang Diklasifikasikan dengan Benar 46 adalah 67.6471%, Instance yang Diklasifikasikan dengan Salah 22 adalah 32.3529 %, Statistik Kappa 0.3725, Berarti kesalahan absolut 0,2993, Root mean kuadrat kesalahan 0,3953, Kesalahan absolut relatif 78,5524 %, Kesalahan kuadrat relatif akar 92,6317 %, Jumlah Total Instance 68 data.

**Tabel 5.13 Detailed Accuracy By Class Percentage split 66% Klasifikasi Native****Bayes Dengan Tools Weka**

<i>TP Rate</i>	<i>FP Rate</i>	<i>Precision</i>	<i>Recall</i>	<i>F-Measure</i>	<i>MCC</i>	<i>ROC Area</i>	<i>PRC Area</i>	<i>Class</i>
0.825	0.393	0.750	0.825	0.786	0.445	0.794	0.867	Neaktif
0.522	0.178	0.600	0.522	0.558	0.357	0.712	0.513	Hipertensi
0.200	0.048	0.663	0.676	0.667	0.395	0.753	0.705	Hipertensi Grad 2
0.676	0.295	0.663	0.676	0.667	0.395	0.753	0.705	<i>Weighted Avg.</i>

Berdasarkan tabel 5.13 rata-rata (*weighted avg.*) pada *detail accuracy by class Percentage split 66%* adalah nilai *TP Rate* adalah 0.676, nilai *FP Rate* adalah 0.295, nilai *Precision* adalah 0.663, nilai *recall* adalah 0.676, nilai *f-measure* adalah 0.667, nilai *mcc* adalah 0.395, nilai *roc area* adalah 0.753 dan nilai *prc area* adalah 0.705.

**Tabel 5.14 Confusion Matrix Percentage split 66% Klasifikasi Native Bayes****Dengan Tools Weka**

<i>Classified as</i>	Negatif	Hipertensi	Hipertensi Grad 2
A	33	5	2
B	10	12	1
C	1	3	1

Berdasarkan tabel 5.14 nilai akhir dari native bayes adalah classified as A untuk negatif 33, hipertensi 5 dan hipertensi grad 2 adalah 2, selanjutnya classified as B untuk negatif 10, hipertensi 12, dan hipertensi grad 2 adalah 1, kemudian classified as C untuk negatif 1, hipertensi 3 dan hipertensi grad 2 adalah 1.

#### 5. Perbandingan *Test Options*

Berikut adalah perbandingan *test options use training set, supplied test set 66%, cros validation folds 5, dan cros validation folds 10.*

**Tabel 5.15 Perbandingan Test Options**

<i>Model Evaluation</i>	Akurasi	Jumlah Kelas	Persentasi
<i>Cros Validation Folds 10</i>	<i>Correctly Classified Instances</i>	110	59%
	<i>Incorrectly Classified Instances</i>	82	41%
<i>Cros Validation Folds 5</i>	<i>Correctly Classified Instances</i>	117	58.5%
	<i>Incorrectly Classified Instances</i>	83	41.5%
<i>Use Training Set</i>	<i>Correctly Classified Instances</i>	130	65%
	<i>Incorrectly Classified Instances</i>	70	35%
<i>Percentage Split 66%</i>	<i>Correctly Classified Instances</i>	46	67.6471%
	<i>Incorrectly Classified Instances</i>	22	32.3529%

Berdasarkan perbandingan *test options* diatas, *cros validation folds 10* memiliki akurasi *correctly classified instances* berjumlah 110 kelas dengan persentasi 59%, dan *incorrectly classified instances* berjumlah 82 kelas dengan

persentasi 41%, selanjutnya *cross validation folds 5* memiliki akurasi *correctly classified instances* berjumlah 117 kelas dengan persentasi 58.5%, dan *incorrectly classified instances* berjumlah 83 kelas dengan persentasi 41.5%, kemudian *use training set* memiliki akurasi *correctly classified instances* berjumlah 130 kelas dengan persentasi 65%, dan *incorrectly classified instances* berjumlah 70 kelas dengan persentasi 35%, dan *Percentage Split 66%* memiliki akurasi *correctly classified instances* berjumlah 46 kelas dengan persentasi 67.6471%, dan *incorrectly classified instances* berjumlah 22 kelas dengan persentasi 32.3529%.

