

BAB V

HASIL ANALISIS DAN VISUALISASI

5.1 HASIL VISUALISASI DATA DENGAN MENGGUNAKAN WEKA

Berikut merupakan bentuk visualisasi menggunakan *tool WEKA* dari beberapa atribut yaitu sebagai berikut :

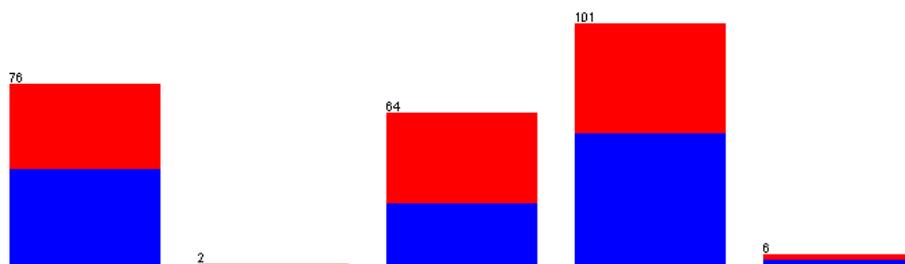
5.1.1 Visualisasi Atribut Usia

Selected attribute

Name: Usia	Distinct: 5	Type: Nominal
Missing: 0 (0%)		Unique: 0 (0%)

No.	Label	Count	Weight
1	>40	76	76.0
2	26	2	2.0
3	<=40	64	64.0
4	26..30	101	101.0
5	<= 25	6	6.0

Class: Kinerja (Nom) Visualize All

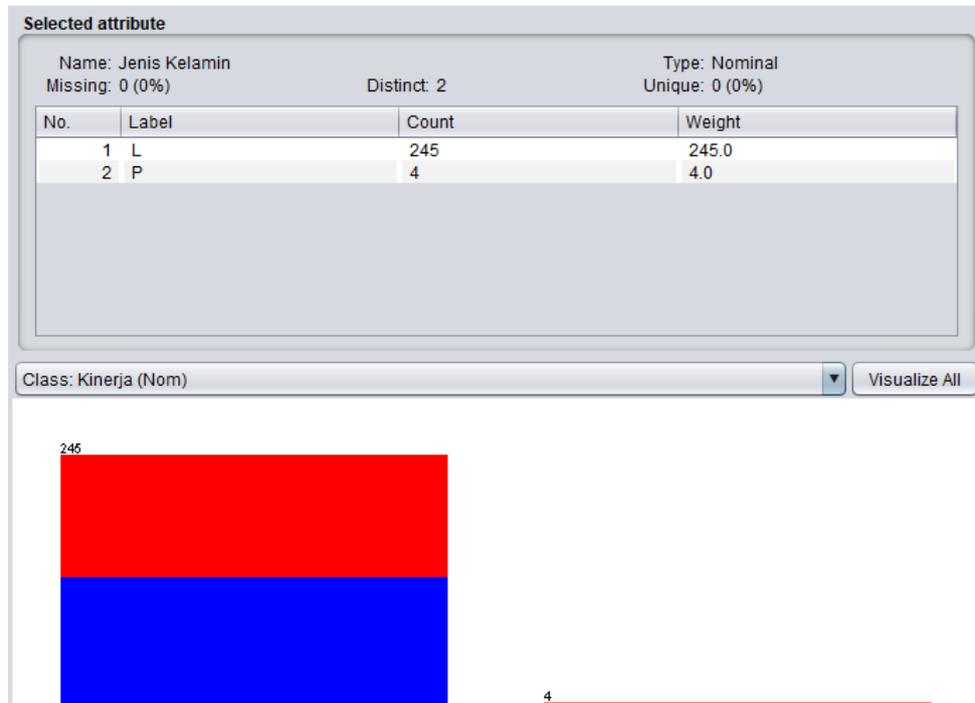


Gambar 5.1 Visualisasi Atribut Usia

Gambar 5.1 adalah visualisasi dari atribut Usia. Diketahui bahwa dari 249 data Satpam di PT. KAJANG LAKO terdapat 76 Satpam yang berumur “>40” dan 2 Satpam yang berumur “26” dan 64 Satpam yang berumur “<=40” dan 101 Satpam yang berumur “26.30” dan 6 Satpam yang berumur “<=25”, maka dapat

disimpulkan bahwa jumlah Satpam yang berumur 26.30 lebih banyak daripada jumlah Satpam yang berumur lainnya.

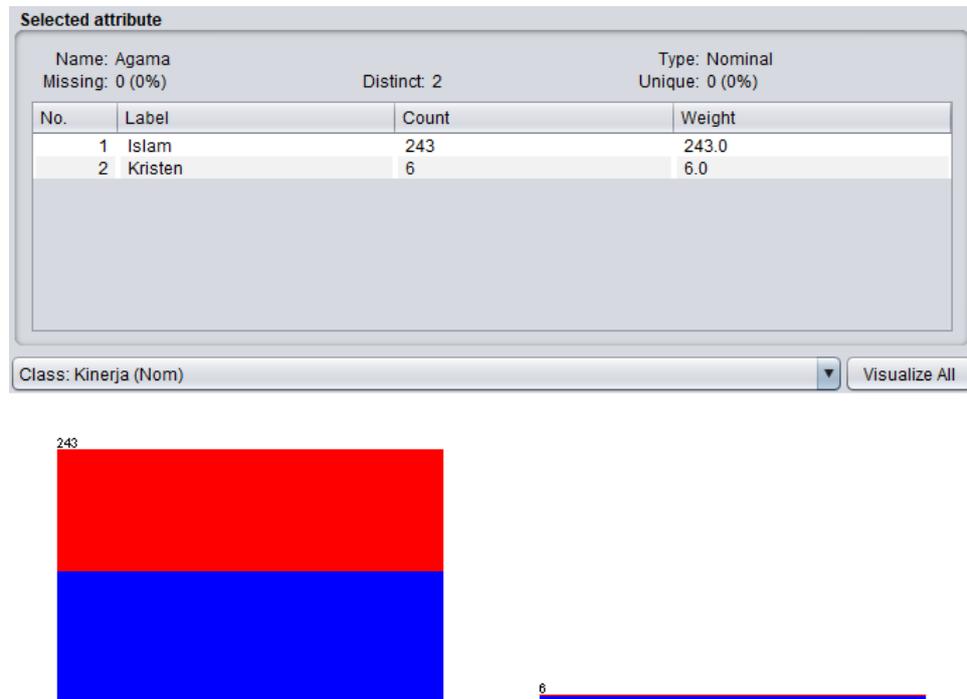
5.1.2 Visualisasi Atribut Jenis Kelamin



Gambar 5.2 Visualisasi Atribut Jenis Kelamin

Gambar 5.2 adalah visualisasi dari atribut Jenis Kelamin. Diketahui bahwa dari 249 data Satpam di PT. KAJANG LAKO terdapat 245 Satpam Berjenis Kelamin “L” dan Terdapat 4 Data Satpam Berjenis Kelamin “P”, maka dapat disimpulkan bahwa jumlah Satpam yang Berjenis Kelamin “L” lebih banyak daripada jumlah Satpam yang berjenis kelamin “P”.

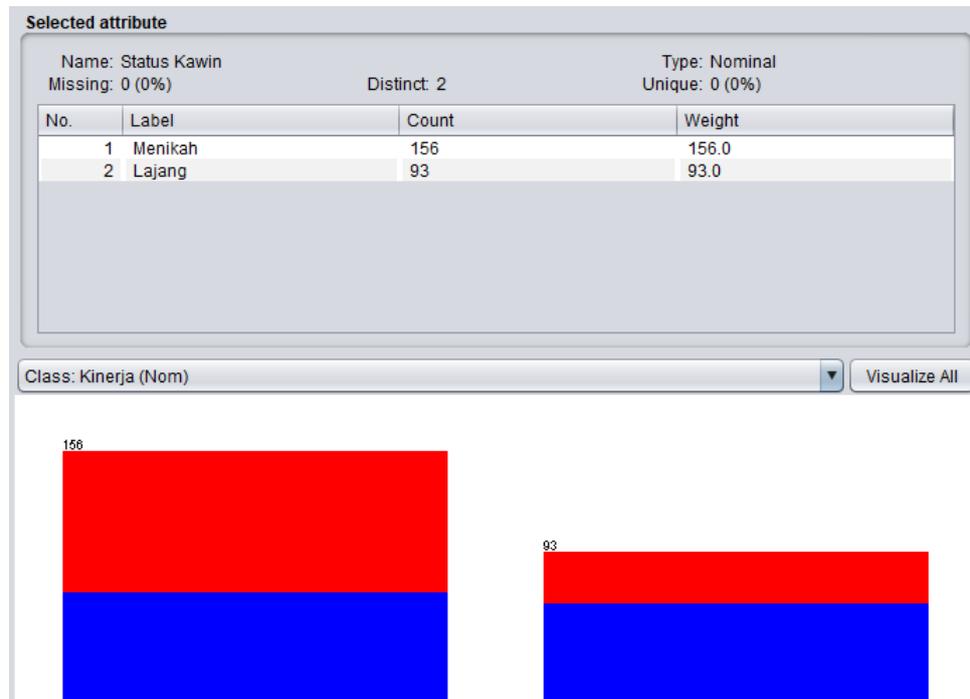
5.1.3 Visualisasi Atribut Agama



Gambar 5.3 Visualisasi Atribut Agama

Gambar 5.3 adalah visualisasi dari atribut Agama. Diketahui bahwa dari 249 data Satpam di PT. KAJANG LAKO terdapat 243 Satpam Ber Agama “ISLAM” dan Terdapat 6 Data Satpam Ber Agama “Kristen”, maka dapat disimpulkan bahwa jumlah Satpam yang Beragama “Islam” lebih banyak dari pada jumlah Satpam yang ber Agama “ISLAM”

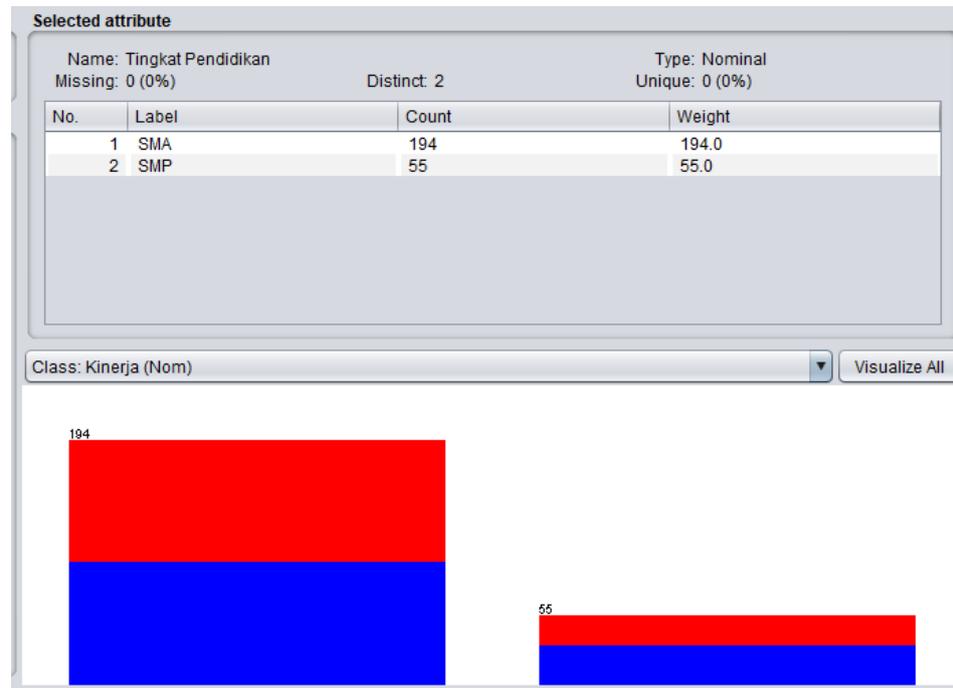
5.1.4 Visualisasi Atribut Status Kawin



Gambar 5.4 Visualisasi Atribut Status Kawin

Gambar 5.4 adalah visualisasi dari atribut Status Kawin. Diketahui bahwa dari 249 data Satpam di PT. KAJANG LAKO terdapat 156 Satpam Status kawin “Menikah” dan Terdapat 93 Data Satpam Status Kawin “Lajang”, maka dapat disimpulkan bahwa jumlah Satpam yang Status Kawin “Menikah” lebih banyak daripada jumlah Satpam yang Status Kawin “Lajang”.

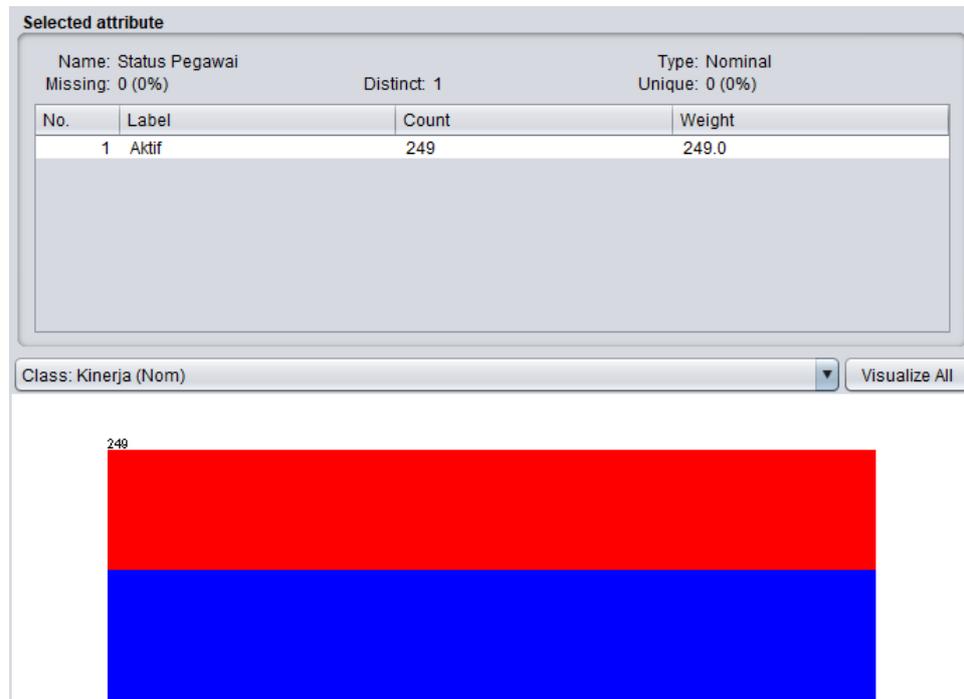
5.1.5 Visualisai Atribut Tingkat Pendidikan



Gambar 5.5 Visualisasi Atribut Tingkat Pendidikan

Gambar 5.5 adalah visualisasi dari atribut Tingkat Pendidikan. Diketahui bahwa dari 249 data Satpam di PT. KAJANG LAKO terdapat 194 Satpam Tingkat Pendidikan “SMA” dan Terdapat 55 Satpam Tingkat Pendidikan “SMP”, maka dapat disimpulkan bahwa jumlah Satpam yang Tingkat Pendidikan “SMA” lebih banyak daripada jumlah Satpam yang Tingkat Pendidikan “SMP”.

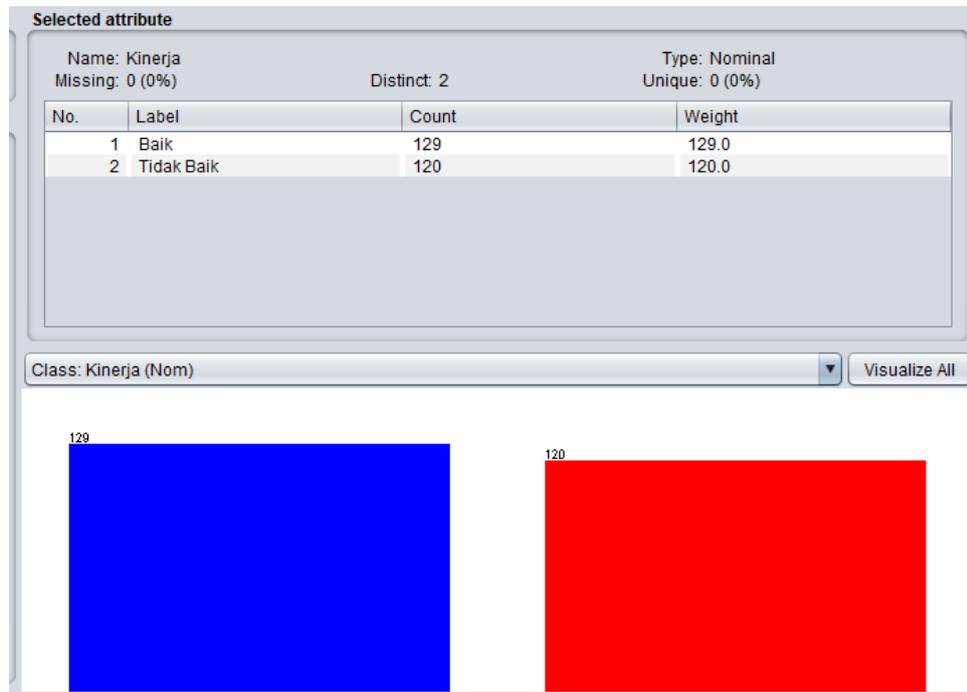
5.1.6 Visualisasi Atribut Status Pegawai



Gambar 5.6 Visualisasi Atribut Status Pegawai

Gambar 5.6 adalah visualisasi dari atribut Status Pegawai. Diketahui bahwa dari 249 data Satpam di PT. KAJANG LAKO terdapat 249 Satpam Status Pegawai “Aktif”, maka dapat disimpulkan bahwa jumlah Satpam yang Status Pegawai “Aktif” Sangat banyak.

5.1.7 Visualisai Atribut Kinerja



Gambar 5.7 Visualisasi Atribut Kinerja

Gambar 5.7 adalah visualisasi dari atribut Kinerja. Diketahui bahwa dari 249 data Satpam di PT. KAJANG LAKO terdapat 129 Satpam Kinerja “Baik” dan Terdapat 120 Satpam Kinerja “Tidak Baik”, maka dapat disimpulkan bahwa jumlah Satpam yang Kinerja “Baik” lebih banyak daripada jumlah Satpam yang Kinerja “Tidak Baik”.

5.2 Hasil Klasifikasi *Naive Bayes* Menggunakan *WEKA*

Klasifikasi status Kinerja Satpam menggunakan algoritma *naive bayes* dilakukan dengan 5 test yaitu menggunakan *Use Training Set*, *5 Cross-Validation*, *10 Cross-Validation*, dan *60% Percentage Split*.

Berikut merupakan hasil klasifikasi *naive bayes* menggunakan *tool WEKA* :

1. Hasil Klasifikasi *Naive Bayes* Menggunakan *Tool WEKA (Use Training Set)*

Test *Use Training Set* melakukan pengujian data menggunakan data *Training* itu sendiri.

The screenshot displays the WEKA interface for the NaiveBayes classifier. The 'Test options' section shows 'Use training set' selected with 10 folds and 66% split. The 'Classifier output' section provides a summary of performance metrics and a detailed accuracy table.

Metric	Value
Correctly Classified Instances	222
Incorrectly Classified Instances	27
Kappa statistic	0.7827
Mean absolute error	0.3143
Root mean squared error	0.336
Relative absolute error	62.937 %
Root relative squared error	67.2394 %
Total Number of Instances	249

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
Weighted Avg.	0.907	0.125	0.886	0.907	0.897	0.783	0.979	0.982	Baik
	0.875	0.093	0.897	0.875	0.886	0.783	0.979	0.978	Tidak Baik


```
==== Confusion Matrix ====
 a  b  <-- classified as
117 12 | a = Baik
 15 105 | b = Tidak Baik
```

Gambar 5.8 Klasifikasi *Naive Bayes (Use Training Set)*

Gambar 5.8 merupakan hasil klasifikasi *naive bayes* pada *tools WEKA* dengan menggunakan *use training set* yang menunjukkan hasil 222 prediksi benar dengan akurasi sebesar 89.1566 % dan 27 prediksi salah dengan persentasi 10.8434% dengan waktu klasifikasi selama 0,06 detik.

2. Hasil Klasifikasi *Naive Bayes* Menggunakan *Tool WEKA* (5 *Cross-Validation*)

Test 5 *Cross-Validation* melakukan pengetesan data dimana data *training* dibagi menjadi k buah *subset* (subhimpunan). 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 menjadi data training sebanyak k-1 kali.

The screenshot displays the WEKA interface for a Naive Bayes classifier. The 'Test options' panel on the left shows 'Cross-validation' selected with 'Folds' set to 5. The 'Classifier output' panel on the right shows the following results:

```

Time taken to build model: 0 seconds
=== Stratified cross-validation ===
=== Summary ===
Correctly Classified Instances      144          57.8313 %
Incorrectly Classified Instances    105          42.1687 %
Kappa statistic                    0.1606
Mean absolute error                 0.4801
Root mean squared error             0.4996
Relative absolute error             96.1489 %
Root relative squared error         99.9843 %
Total Number of Instances          249

=== Detailed Accuracy By Class ===
              TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
              0.512   0.350   0.611     0.512   0.557     0.163   0.584    0.562    Baik
              0.650   0.488   0.553     0.650   0.598     0.163   0.584    0.556    Tidak Baik
Weighted Avg.   0.578   0.417   0.583     0.578   0.577     0.163   0.584    0.559

=== Confusion Matrix ===
  a  b  <-- classified as
66 63 | a = Baik
42 78 | b = Tidak Baik

```

Gambar 5.9 Klasifikasi *Naive Bayes* (5 *Cross-Validation*)

Gambar 5.9 merupakan hasil klasifikasi *naive bayes* pada *tools WEKA* dengan menggunakan 5 *Cross-Validation* yang menunjukkan hasil 144 prediksi benar dengan akurasi sebesar 57.8313% dan 105 prediksi salah dengan persentasi 42.1687% dengan waktu klasifikasi selama 0 detik.

3. Hasil Klasifikasi *Naive Bayes* Menggunakan *Tool WEKA* (10 *Cross-Validation*)

Test 10 *Cross-Validation* melakukan pengetesan data dimana 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 menjadi data *Training* Sebanyak *k-1* kali

The screenshot displays the WEKA interface for a Naive Bayes classifier. The 'Test options' section shows 'Cross-validation' selected with 'Folds' set to 10. The 'Classifier output' section provides a summary of performance metrics and a detailed accuracy table by class.

Classifier output Summary:

```

Time taken to build model: 0 seconds
=== Stratified cross-validation ===
=== Summary ===
Correctly Classified Instances      138           55.4217 %
Incorrectly Classified Instances    111           44.5783 %
Kappa statistic                    0.1132
Mean absolute error                0.4796
Root mean squared error            0.4989
Relative absolute error            96.0467 %
Root relative squared error        99.8534 %
Total Number of Instances          249
  
```

Detailed Accuracy By Class:

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.481	0.367	0.585	0.481	0.528	0.115	0.585	0.563	Baik
	0.633	0.519	0.531	0.633	0.578	0.115	0.585	0.556	Tidak Baik
Weighted Avg.	0.554	0.440	0.559	0.554	0.552	0.115	0.585	0.559	

Confusion Matrix:

```

=== Confusion Matrix ===
 a b  <-- classified as
62 67 | a = Baik
44 76 | b = Tidak Baik
  
```

The 'Result list' on the left shows three runs of the classifier, with the most recent run (21:12:25) selected.

Gambar 5.10 Klasifikasi *Naive Bayes* (10 *Cross-Validation*)

Gambar 5.10 merupakan hasil klasifikasi *naive bayes* pada *tools* WEKA dengan menggunakan 10 *Cross-Validation* yang menunjukkan hasil 138 prediksi benar dengan akurasi sebesar 55.4217% dan 111 prediksi salah dengan persentasi 44.5783% dengan waktu klasifikasi selama 0 detik.

4. Hasil Klasifikasi *Naive Bayes* Menggunakan *Tool WEKA* (60% *Percentage Split*)

Tes *Percentage Split* hasil klasifikasi akan dites dengan menggunakan k% dari data tersebut. Pada tes ini akan digunakan 60% *Percentage Split* dari data.

The screenshot shows the WEKA interface for a Naive Bayes classifier. The 'Test options' panel on the left is configured with 'Percentage split' at 60%. The 'Classifier output' panel on the right displays the following results:

```

=== Evaluation on test split ===
Time taken to test model on test split: 0 seconds

=== Summary ===
Correctly Classified Instances      51      51 %
Incorrectly Classified Instances    49      49 %
Kappa statistic                    0.0452
Mean absolute error                0.4832
Root mean squared error            0.501
Relative absolute error            95.0583 %
Root relative squared error        97.66 %
Total Number of Instances          100

=== Detailed Accuracy By Class ===
               TP Rate  FP Rate  Precision  Recall  F-Measure  MCC   ROC Area  PRC Area  Class
0.636   0.589   0.459   0.636   0.533   0.048   0.579   0.568   Baik
0.411   0.364   0.590   0.411   0.484   0.048   0.579   0.613   Tidak Baik
Weighted Avg.   0.510   0.463   0.532   0.510   0.506   0.048   0.579   0.593

=== Confusion Matrix ===
 a b  <-- classified as
28 16 | a = Baik
33 23 | b = Tidak Baik

```

Gambar 5.11 Klasifikasi *Naive Bayes* (60% *Percentage Split*)

Gambar 5.11 merupakan hasil klasifikasi *naive bayes* pada *tools* WEKA dengan menggunakan 60% *Percentage Split* yang menunjukkan hasil 51 prediksi benar dengan akurasi sebesar 51% dan 49 prediksi salah dengan persentasi 49% dengan waktu klasifikasi selama 0 detik.

5.4 Hasil Perbandingan Evaluasi Akurasi Dari 5 Test Options

Setelah dilakukan analisis klasifikasi *naive bayes* pada tool WEKA menggunakan *Use Training Set*, *5 Fold Cross Validation*, *10 Fold Cross Validation*, *60% Percentage Split*, dan *80% Percentage Split*, maka diperoleh akurasi tertinggi yaitu dengan menggunakan *Use Training Set* dengan persentasi akurasi yaitu 98.01 % untuk *Correctly Classified Instances* dan 1.99 % untuk *Incorrectly Classified Instances*. Perbandingan hasil analisis dapat dilihat pada tabel 5.1.

Tabel 5.1 Perbandingan Evaluasi Akurasi WEKA

Model Evaluasi	Akurasi	Jumlah Kelas	Persentasi
<i>Use Training Set</i>	<i>Correctly Classified Instances</i>	222	89.1566 %
	<i>Incorrectly Classified Instances</i>	27	10.8434%
<i>5 Fold Cross Validation</i>	<i>Correctly Classified Instances</i>	144	57.8313%
	<i>Incorrectly Classified Instances</i>	105	42.1687%
<i>10 Fold Cross Validation</i>	<i>Correctly Classified Instances</i>	139	55.4217%
	<i>Incorrectly Classified Instances</i>	111	44.5783%
<i>60% Percentage Split</i>	<i>Correctly Classified Instances</i>	51	51 %
	<i>Incorrectly Classified Instances</i>	49	49 %

Persentasi Hasil Klasifikasi *Naive Bayes* dengan menggunakan 4 *test options*. Pada *Use Training Set* dengan jumlah kelas *Correctly Classified Instances* 222, *Incorrectly Classified Instances* 27, dan persentasi akurasi *Correctly Classified Instances* sebesar **89.1566 %**, *Incorrectly Classified Instances* **10.8434%**. Pada tes 5 *Fold Cross Validation* dengan jumlah kelas *Classified Instances* 144, *Incorrectly Classified Instances* 105, dan persentasi akurasi *Correctly Classified Instances* sebesar **57.8313%**, *Incorrectly Classified Instances* **42.1687%**. Pada tes 10 *Fold Cross Validation* dengan jumlah kelas *Correctly Classified Instances* 139, *Incorrectly Classified Instances* 111, dan persentasi akurasi *Correctly Classified Instances* sebesar **55.4217%** , *Incorrectly Classified Instances* **44.5783%**. Pada tes 60% *Percentage Split* dengan jumlah kelas *Correctly Classified Instances* 51, *Incorrectly Classified Instances* 49, dan persentasi akurasi *Correctly Classified Instances* sebesar 51%, *Incorrectly Classified Instances* 49%.