

## BAB V

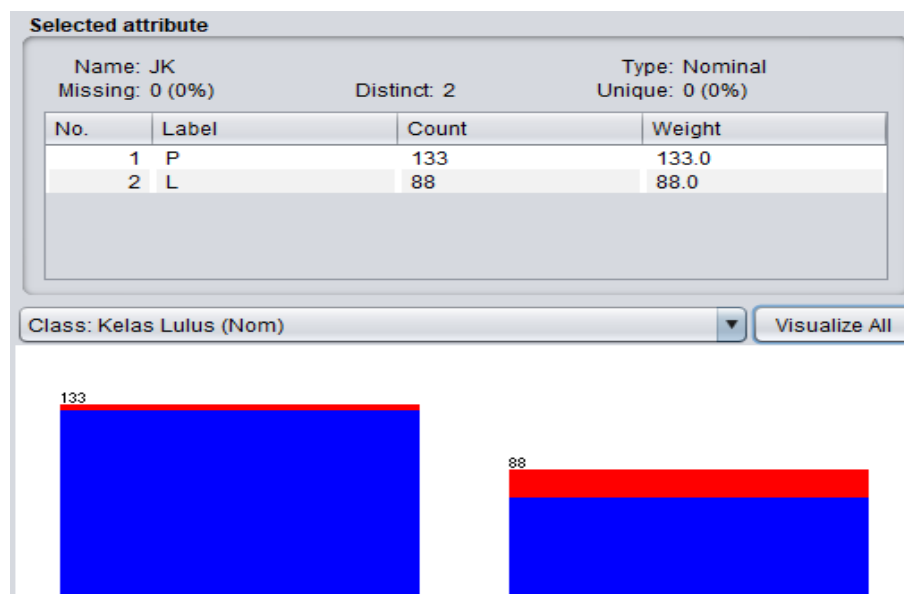
### HASIL ANALISIS DAN VISUALISASI

#### 5.1 HASIL VISUALISASI DATA DENGAN MENGGUNAKAN WEKA

##### 5.1.1 Hasil Visualisasi Atribut Dengan Menggunakan *Tool WEKA*

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

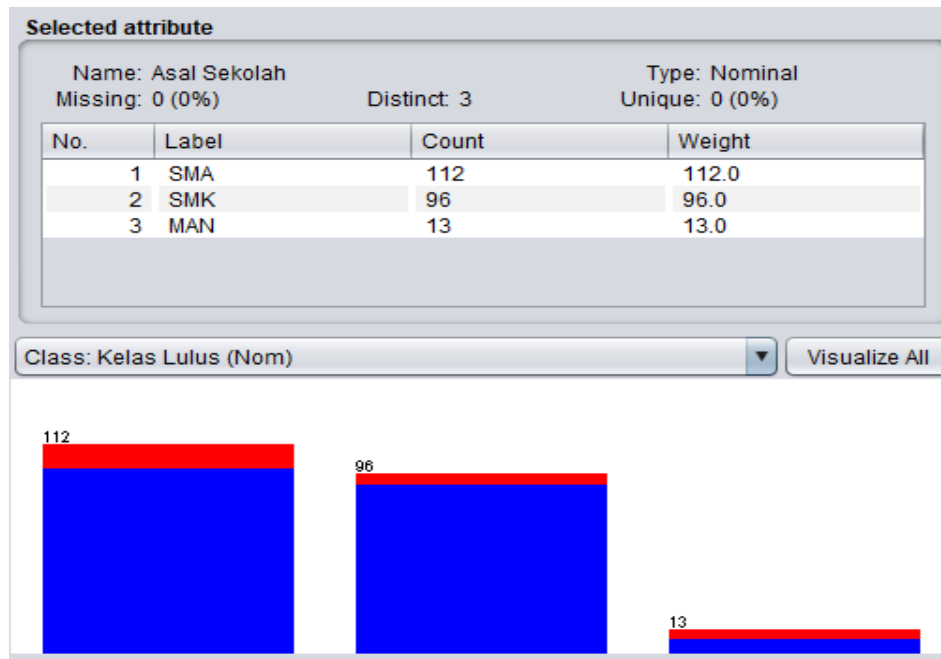
1. Visualisasi Atribut Jenis Kelamin



**Gambar 5.1 Visualisasi Atribut Jenis Kelamin**

Gambar 5.1 adalah visualisasi dari atribut Jenis Kelamin. Diketahui bahwa dari 221 data mahasiswa Sistem Informasi terdapat 133 mahasiswa yang berjenis kelamin “L (Laki-laki)”, dan 88 mahasiswa yang berjenis kelamin “P (Perempuan)”, maka dapat disimpulkan bahwa jumlah mahasiswa yang berjenis kelamin laki-laki lebih banyak daripada jumlah mahasiswa yang berjenis kelamin perempuan.

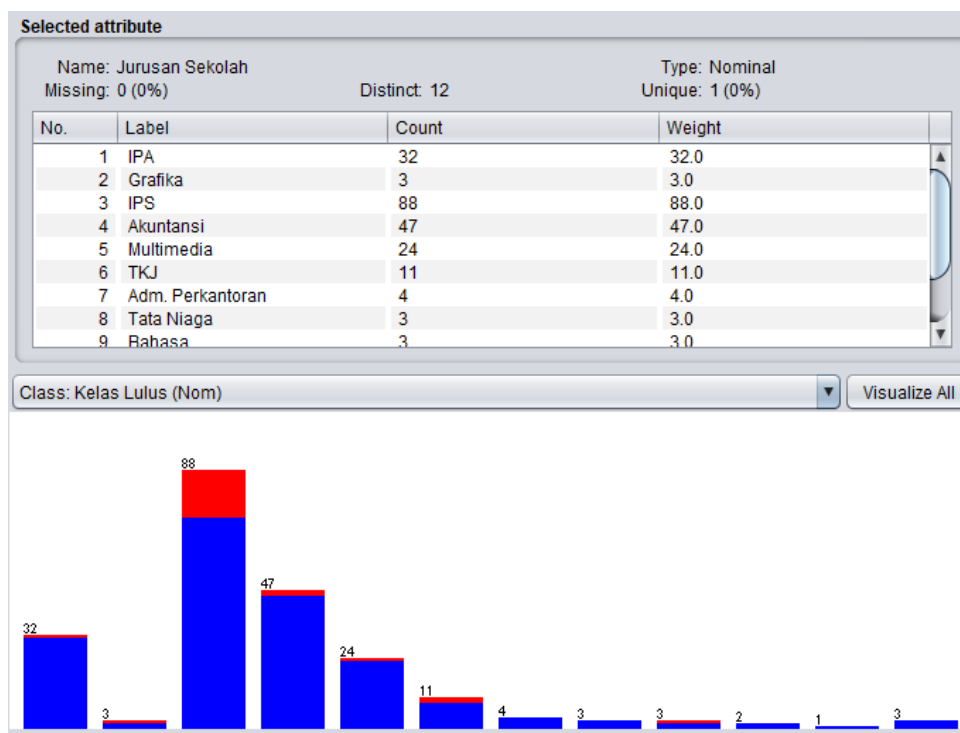
## 2. Visualisasi Atribut Asal Sekolah



**Gambar 5.2 Visualisasi Atribut Asal Sekolah**

Gambar 5.2 adalah visualisasi dari atribut Asal Sekolah. Diketahui bahwa dari 221 data mahasiswa Sistem Informasi terdapat 112 mahasiswa yang berasal dari “SMA”, 96 mahasiswa yang berasal dari “SMK”, dan 13 mahasiswa yang berasal dari “MAN”. Maka dapat disimpulkan bahwa jumlah mahasiswa dengan Asal Sekolah “SMA” lebih banyak daripada mahasiswa dengan Asal Sekolah “SMK” dan “MAN”.

### 3. Visualisasi Atribut Jurusan Sekolah

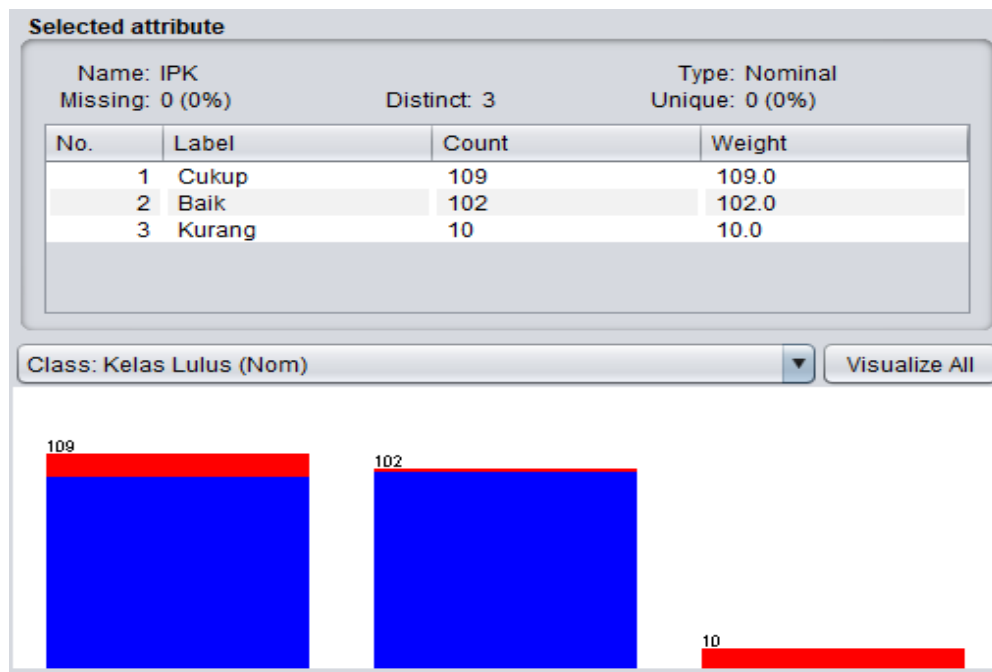


**Gambar 5.3 Visualisasi Atribut Jurusan Sekolah**

Gambar 5.3 adalah visualisasi dari atribut Jurusan Sekolah. Diketahui bahwa dari 221 data mahasiswa Sistem Informasi terdapat 32 mahasiswa dengan Jurusan Sekolah “IPA”, 3 mahasiswa dengan Jurusan Sekolah “Grafika”, 88 mahasiswa dengan Jurusan Sekolah “IPS”, 47 mahasiswa dengan Jurusan Sekolah “Akuntansi”, 24 mahasiswa dengan Jurusan Sekolah “Multimedia”, 11 mahasiswa dengan Jurusan Sekolah “TKJ”, 4 mahasiswa dengan Jurusan Sekolah “Adm. Perkantoran”, 3 mahasiswa dengan Jurusan Sekolah “Tata Niaga”, 3 mahasiswa dengan Jurusan Sekolah “Bahasa”, 2 mahasiswa dengan Jurusan Sekolah “Agama”, 1 mahasiswa dengan Jurusan Sekolah “Busana Butik”, dan 3 mahasiswa dengan Jurusan Sekolah “Akomodasi Perhotelan”. Maka dapat

disimpulkan bahwa jumlah mahasiswa dengan jurusan sekolah “IPS” lebih banyak daripada jumlah mahasiswa dengan jurusan sekolah lainnya.

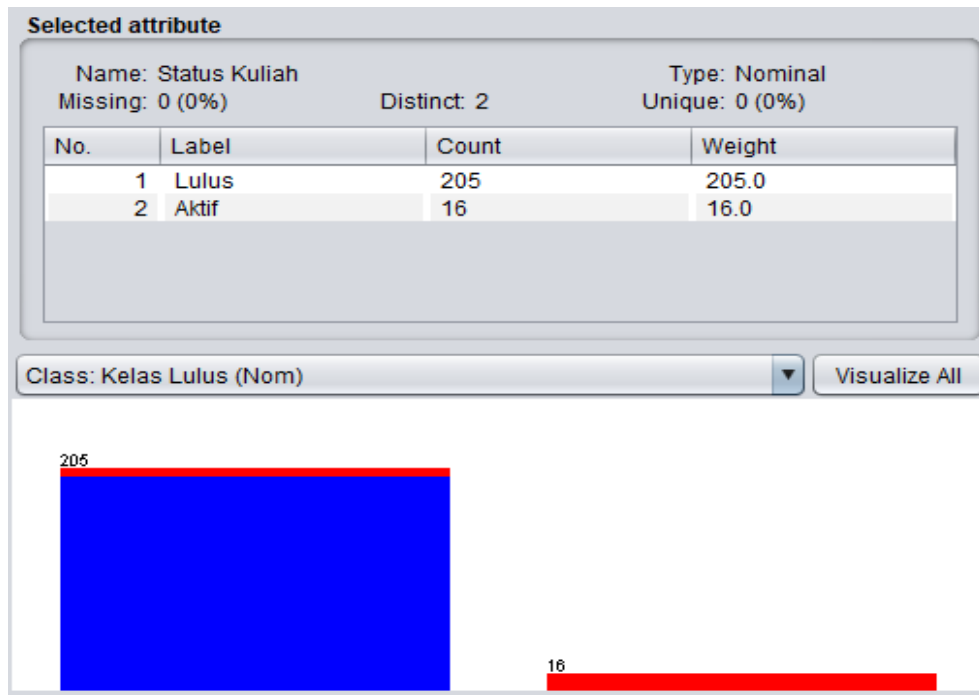
#### 4. Visualisasi Atribut IPK



**Gambar 5.4 Visualisasi Atribut IPK**

Gambar 5.4 adalah visualisasi dari atribut IPK. Diketahui bahwa dari 221 data mahasiswa Sistem Informasi terdapat 109 mahasiswa yang memiliki IPK “Cukup”, 102 mahasiswa yang memiliki IPK “Baik”, dan 10 mahasiswa yang mempunyai IPK “Kurang”. Maka dapat disimpulkan bahwa jumlah mahasiswa dengan IPK “Cukup” lebih banyak daripada jumlah mahasiswa dengan IPK “Baik” dan “Kurang”.

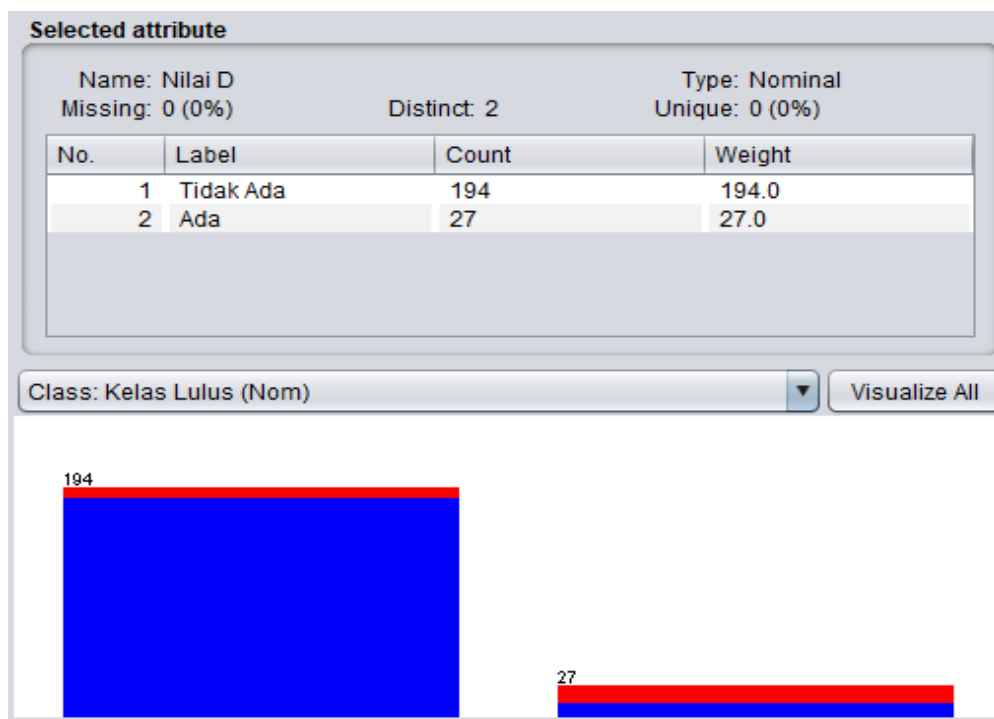
## 5. Visualisasi Atribut Status Kuliah



**Gambar 5.5 Visualisasi Atribut Status Kuliah**

Gambar 5.5 adalah visualisasi dari atribut Status Kuliah. Diketahui bahwa dari 221 data mahasiswa Sistem Informasi terdapat 205 mahasiswa dengan Status Kuliah “Lulus”, dan 16 mahasiswa dengan Status Kuliah “Aktif”. Maka dapat disimpulkan bahwa jumlah mahasiswa dengan Status Kuliah “Lulus” lebih banyak daripada jumlah mahasiswa dengan Status Kuliah “Aktif”.

## 6. Visualisasi Atribut Nilai D



**Gambar 5.6 Visualisasi Atribut Nilai D**

Gambar 5.6 adalah visualisasi dari atribut Nilai D. Diketahui bahwa dari 221 data mahasiswa Sistem Informasi terdapat 194 mahasiswa yang memiliki Nilai D “Tidak Ada”, dan 27 mahasiswa dengan Nilai D “Ada”. Maka dapat disimpulkan bahwa jumlah mahasiswa dengan Nilai D “Tidak Ada” lebih banyak daripada jumlah mahasiswa dengan Nilai D “Ada”.

### 5.1.2 Hasil Klasifikasi Algoritma C4.5 Menggunakan Tool WEKA

Klasifikasi kelas lulus mahasiswa Sistem Informasi menggunakan algoritma algoritma C4.5 dilakukan dengan 5 test yaitu menggunakan *Use Data Training*, *5 Cross-Validation*, *10 Cross-Validation*, *60% Percentage Split*, dan

80% *Percentage Split*. Berikut merupakan hasil klasifikasi *naive bayes* menggunakan *tool WEKA* :

1. Hasil Klasifikasi Algoritma C4.5 Menggunakan *Tool WEKA (Use Data Training)*

Test *Use Data Training* melakukan pengetesan data menggunakan data *training* itu sendiri.

The screenshot shows the WEKA interface with the following details:

- Classifier:** J48 -C 0.25-M 2
- Test options:**
  - Use training set
  - Supplied test set (Set...)
  - Cross-validation (Folds: 10)
  - Percentage split (%: 66)
  - More options...
- Classifier output:**

```

=== Evaluation on training set ===

Time taken to test model on training data: 0.02 seconds

=== Summary ===

Correctly Classified Instances      213      96.3801 %
Incorrectly Classified Instances     8        3.6199 %
Kappa statistic                     0.781
Mean absolute error                  0.0696
Root mean squared error              0.1865
Relative absolute error              35.4321 %
Root relative squared error          59.942 %
Total Number of Instances           221

=== Detailed Accuracy By Class ===

          TP Rate  FP Rate  Precision  Recall  F-Measure  MCC   ROC Area  PRC Area  Class
          -----  -
          1.000   0.333   0.961     1.000   0.980     0.800  0.833   0.961   Tepat
          0.667   0.000   1.000     0.667   0.800     0.800  0.833   0.703   Lambat
Weighted Avg.   0.964   0.297   0.965     0.964   0.961     0.800  0.833   0.933

=== Confusion Matrix ===

 a  b  <-- classified as
197  0  |  a = Tepat
 8 16  |  b = Lambat

```
- Result list:** 21:39:31 - trees.J48

**Gambar 5.7** Klasifikasi Algoritma C4.5 (*Use Data Training*)

Gambar 5.7 merupakan hasil klasifikasi algoritma C4.5 pada *tools WEKA* dengan menggunakan *use training set* yang menunjukkan hasil 213 prediksi benar

dengan akurasi sebesar 96.3801% dan 8 prediksi salah dengan persentasi 3.6199% dengan waktu klasifikasi selama 0.02 detik.

## 2. Hasil Klasifikasi Algoritma C4.5 Menggunakan *Tool WEKA (5 Cross-Validation)*

The screenshot displays the WEKA interface for a C4.5 classifier. The 'Test options' section shows 'Cross-validation' selected with 'Folds' set to 5. The 'Classifier output' section provides a summary of performance metrics and a detailed accuracy breakdown by class.

**Classifier output Summary:**

Metric	Value	Percentage
Correctly Classified Instances	213	96.3801 %
Incorrectly Classified Instances	8	3.6199 %
Kappa statistic	0.781	
Mean absolute error	0.0698	
Root mean squared error	0.1882	
Relative absolute error	35.4072 %	
Root relative squared error	60.4749 %	
Total Number of Instances	221	

**Detailed Accuracy By Class:**

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	1.000	0.333	0.961	1.000	0.980	0.800	0.750	0.941	Tepat
	0.667	0.000	1.000	0.667	0.800	0.800	0.750	0.703	Lambat
Weighted Avg.	0.964	0.297	0.965	0.964	0.961	0.800	0.750	0.915	

**Confusion Matrix:**

```

=== Confusion Matrix ===
 a  b  <-- classified as
197  0  | a = Tepat
  8 16 | b = Lambat

```

The 'Result list' on the left shows two entries: '21:39:31 - trees.J48' and '21:42:45 - trees.J48', with the second entry selected.

**Gambar 5.8 Klasifikasi Algoritma C4.5 (5 Cross-Validation)**

Gambar 5.8 merupakan hasil klasifikasi algoritma C4.5 pada *tools WEKA* dengan menggunakan 5 *Cross-Validation* yang menunjukkan hasil 213 prediksi benar dengan akurasi sebesar 96.3801% dan 8 prediksi salah dengan persentasi 3.6199% dengan waktu klasifikasi selama 0 detik.



### 3. Hasil Klasifikasi Algoritma C4.5 Menggunakan *Tool WEKA (10 Cross-Validation)*

The screenshot shows the WEKA interface with the following details:

- Test options:**
  - Use training set:
  - Supplied test set:  (Set...)
  - Cross-validation:  Folds: 10
  - Percentage split:  % 66
  - More options... button
  - Class: (Nom) Kelas Lulus
  - Start and Stop buttons
- Classifier output:**
  - Time taken to build model: 0.01 seconds
  - === Stratified cross-validation ===
  - === Summary ===
  - Correctly Classified Instances: 213 (96.3801 %)
  - Incorrectly Classified Instances: 8 (3.6199 %)
  - Kappa statistic: 0.781
  - Mean absolute error: 0.0698
  - Root mean squared error: 0.1875
  - Relative absolute error: 35.4525 %
  - Root relative squared error: 60.2236 %
  - Total Number of Instances: 221
  - === Detailed Accuracy By Class ===
  - Table with columns: TP Rate, FP Rate, Precision, Recall, F-Measure, MCC, ROC Area, PRC Area, Class
  - Weighted Avg. row: 0.964, 0.297, 0.965, 0.964, 0.961, 0.800, 0.746, 0.912
  - === Confusion Matrix ===
  - Confusion matrix table:

a	b	<-- classified as
197	0	a = Tepat
8	16	b = Lambat

**Gambar 5.9** Klasifikasi Algoritma C4.5 (*10 Cross-Validation*)

Gambar 5.9 merupakan hasil klasifikasi algoritma C4.5 pada *tools WEKA* dengan menggunakan *10 Cross-Validation* yang menunjukkan hasil 213 prediksi benar dengan akurasi sebesar 96.3801% dan 8 prediksi salah dengan persentasi 3.6199% dengan waktu klasifikasi selama 0.01 detik.

#### 4. Hasil Klasifikasi Algoritma C4.5 Menggunakan Tool WEKA (70% Percentage Split)

Tes *Percentage Split* hasil klasifikasi akan dites dengan menggunakan k% dari data tersebut. Pada tes ini akan digunakan 70% *Percentage Split* dari data.

The screenshot shows the WEKA interface with the following details:

- Test options:**
  - Use training set:
  - Supplied test set:  Set..
  - Cross-validation:  Folds: 10
  - Percentage split:  % 70
  - More options... button
  - (Nom) Kelas Lulus dropdown
  - Start and Stop buttons
- Classifier output:**

```

=== Evaluation on test split ===

Time taken to test model on test split: 0 seconds

=== Summary ===

Correctly Classified Instances      62      93.9394 %
Incorrectly Classified Instances    4       6.0606 %
Kappa statistic                    0.3161
Mean absolute error                 0.0853
Root mean squared error             0.2407
Relative absolute error             46.3877 %
Root relative squared error        89.2884 %
Total Number of Instances          66

=== Detailed Accuracy By Class ===

              TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
              1.000   0.000   0.938     1.000   0.968     0.433   0.600    0.938    Tepat
              0.200   0.000   1.000     0.200   0.333     0.433   0.600    0.261    Lambat
Weighted Avg.   0.939   0.739   0.943     0.939   0.920     0.433   0.600    0.887

=== Confusion Matrix ===

 a b  <-- classified as
61  0 | a = Tepat
 4  1 | b = Lambat

```
- Result list (right-click for options):**
  - 21:37:33 - trees.J48
  - 21:37:40 - trees.J48
  - 21:37:44 - trees.J48
  - 21:37:51 - trees.J48 (selected)

**Gambar 5.10 Klasifikasi Algoritma C4.5 (70% Percentage Split)**

Gambar 5.10 merupakan hasil klasifikasi algoritma C4.5 pada tools WEKA dengan menggunakan 70% *Percentage Split* yang menunjukkan hasil 62 prediksi benar dengan akurasi sebesar 93.9394% dan 4 prediksi salah dengan persentasi 6.0606% dengan waktu klasifikasi selama 0 detik.

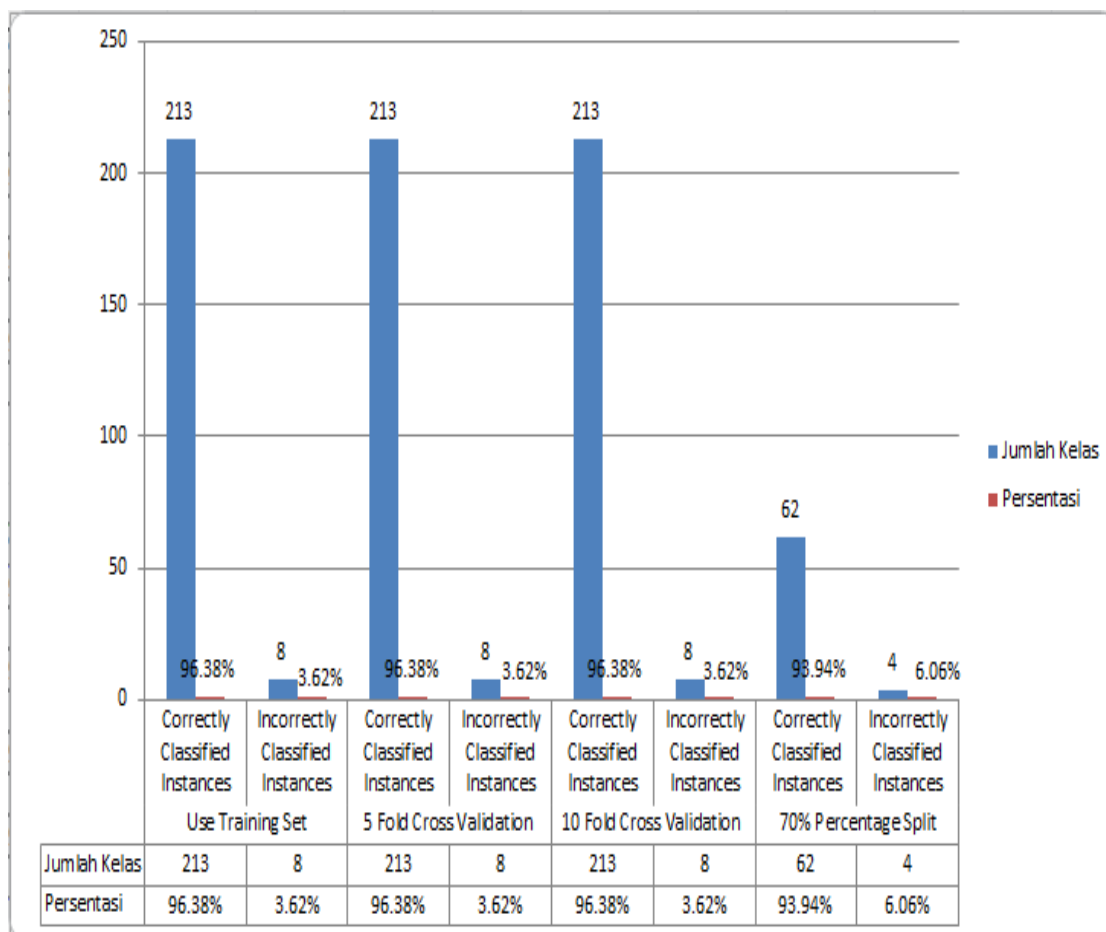
### 5.1.3 Hasil Perbandingan Evaluasi Akurasi Algoritma C4.5 Dari Data Kelulusan Mahasiswa Sistem Informasi

Setelah dilakukan analisis klasifikasi algoritma C4.5 pada *tool WEKA* menggunakan *Use Training Set*, *5 Fold Cross Validation*, *10 Fold Cross Validation*, dan *70% Percentage Split*, maka diperoleh akurasi tertinggi yaitu dengan menggunakan *Use Training Set* dengan persentasi akurasi yaitu 95.4751% untuk *Correctly Classified Instances* dan 4.5249% untuk *Incorrectly Classified Instances*. Perbandingan hasil analisis dapat dilihat pada tabel 5.1.

**Tabel 5.1 Perbandingan Evaluasi Akurasi Algoritma C4.5**

Model Evaluasi	Akurasi	Jumlah Kelas	Persentasi
<i>Use Training Set</i>	<i>Correctly Classified Instances</i>	213	96.3801%
	<i>Incorrectly Classified Instances</i>	8	3.6199%
<i>5 Fold Cross Validation</i>	<i>Correctly Classified Instances</i>	213	96.3801%
	<i>Incorrectly Classified Instances</i>	8	3.6199%
<i>10 Fold Cross Validation</i>	<i>Correctly Classified Instances</i>	213	96.3801%
	<i>Incorrectly Classified Instances</i>	8	3.6199%
<i>70% Percentage Split</i>	<i>Correctly Classified Instances</i>	62	93.9394%
	<i>Incorrectly Classified Instances</i>	4	6.0606%

Grafik hasil akurasi klasifikasi *naïve bayes* menggunakan *tool WEKA* dapat dilihat pada gambar 5.11 :



**Gambar 5.11 Hasil Akurasi Klasifikasi Dengan Algoritma C4.5**

Gambar 5.11 merupakan grafik persentasi hasil klasifikasi algoritma C4.5 dengan menggunakan 4 *test options*. Pada *Use Training Set* dengan jumlah kelas *Correctly Classified Instances* 213, *Incorrectly Classified Instances* 8, dan persentasi akurasi *Correctly Classified Instances* sebesar 96.3801%, *Incorrectly Classified Instances* 3.6199%. Pada tes *5 Fold Cross Validation* dengan jumlah kelas *Classified Instances* 213, *Incorrectly Classified Instances* 8, dan persentasi

akurasi *Correctly Classified Instances* sebesar 96.3801%, *Incorrectly Classified Instances* 3.6199%. Pada tes 10 *Fold Cross Validation* dengan jumlah kelas *Correctly Classified Instances* 213, *Incorrectly Classified Instances* 8, dan persentasi akurasi *Correctly Classified Instances* sebesar 96.3801%, *Incorrectly Classified Instances* 3.6199%. Pada tes 70% *Percentage Split* dengan jumlah kelas *Correctly Classified Instances* 62, *Incorrectly Classified Instances* 4, dan persentasi akurasi *Correctly Classified Instances* sebesar 93.9394%, *Incorrectly Classified Instances* 6.0606%.

#### **5.1.4 Hasil Klasifikasi *Naïve Bayes* Menggunakan *Tool WEKA***

Klasifikasi kelas lulus mahasiswa Sistem Informasi menggunakan algoritma *naïve bayes* dilakukan dengan 4 test yaitu menggunakan *Use Data Training*, 5 *Cross-Validation*, 10 *Cross-Validation*, dan 70% *Percentage Split*. Berikut merupakan hasil klasifikasi *naive bayes* menggunakan *tool WEKA* :

1. Hasil Klasifikasi *Naïve Bayes* Menggunakan *Tool WEKA* (*Use Data Training*)

Test *Use Data Training* melakukan pengetesan data menggunakan data *training* itu sendiri.

Choose NaiveBayes

**Test options**

Use training set  
 Supplied test set Set...  
 Cross-validation Folds 10  
 Percentage split % 66  
 More options...

(Nom) Kelas Lulus

Start Stop

**Result list (right-click for options)**

19:32:49 - bayes.NaiveBayes

**Classifier output**

```

=== Evaluation on training set ===

Time taken to test model on training data: 0.02 seconds

=== Summary ===

Correctly Classified Instances      211          95.4751 %
Incorrectly Classified Instances    10           4.5249 %
Kappa statistic                    0.7478
Mean absolute error                 0.0537
Root mean squared error             0.1777
Relative absolute error             27.3328 %
Root relative squared error         57.1017 %
Total Number of Instances          221

=== Detailed Accuracy By Class ===

          TP Rate  FP Rate  Precision  Recall  F-Measure  MCC   ROC Area  PRC Area  Class
          0.985   0.292   0.965     0.985   0.975     0.752  0.968    0.995    Tepat
          0.708   0.015   0.850     0.708   0.773     0.752  0.968    0.882    Lambat
Weighted Avg.   0.955   0.262   0.953     0.955   0.953     0.752  0.968    0.983

=== Confusion Matrix ===

  a  b  <-- classified as
194 3 | a = Tepat
 7 17 | b = Lambat

```

**Gambar 5.12 Klasifikasi *Naïve Bayes* (Use Data Training)**

Gambar 5.12 merupakan hasil klasifikasi *naïve bayes* pada *tools* WEKA dengan menggunakan *use training set* yang menunjukkan hasil 211 prediksi benar dengan akurasi sebesar 95.4751% dan 10 prediksi salah dengan persentasi 4.5249% dengan waktu klasifikasi selama 0.02 detik.

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

The screenshot displays the WEKA interface for a Naive Bayes classifier. The 'Test options' panel shows 'Cross-validation' selected with 5 folds. The 'Classifier output' panel shows the following summary:

```

Time taken to build model: 0 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      210      95.0226 %
Incorrectly Classified Instances    11       4.9774 %
Kappa statistic                    0.717
Mean absolute error                 0.0635
Root mean squared error             0.1995
Relative absolute error             32.2066 %
Root relative squared error         64.0915 %
Total Number of Instances          221
  
```

The 'Detailed Accuracy By Class' table is as follows:

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.985	0.333	0.960	0.985	0.972	0.723	0.941	0.985	Tepat
	0.667	0.015	0.842	0.667	0.744	0.723	0.941	0.846	Lambat
Weighted Avg.	0.950	0.299	0.948	0.950	0.948	0.723	0.941	0.970	

The 'Confusion Matrix' is shown below:

```

=== Confusion Matrix ===

 a  b  <-- classified as
194  3 | a = Tepat
  8 16 | b = Lambat
  
```

The 'Result list' shows two entries for 'b-yes.NaiveBayes' with timestamps 19:32:49 and 19:57:11.

**Gambar 5.13** Klasifikasi *Naïve Bayes* (5 *Cross-Validation*)

Gambar 5.13 merupakan hasil klasifikasi *naïve bayes* pada *tools WEKA* dengan menggunakan 5 *Cross-Validation* yang menunjukkan hasil 210 prediksi benar dengan akurasi sebesar 95.0226% dan 11 prediksi salah dengan persentase 4.9774% dengan waktu klasifikasi selama 0 detik.

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

The screenshot displays the WEKA interface for a Naive Bayes classifier. The 'Test options' panel on the left shows 'Cross-validation' selected with 10 folds. The 'Classifier output' panel on the right provides the following data:

Time taken to build model: 0 seconds

=== Stratified cross-validation ===  
 === Summary ===

Correctly Classified Instances	210	95.0226 %
Incorrectly Classified Instances	11	4.9774 %
Kappa statistic	0.717	
Mean absolute error	0.06	
Root mean squared error	0.1924	
Relative absolute error	30.4799 %	
Root relative squared error	61.7847 %	
Total Number of Instances	221	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.985	0.333	0.960	0.985	0.972	0.723	0.948	0.989	Tepat
	0.667	0.015	0.842	0.667	0.744	0.723	0.948	0.860	Lambat
Weighted Avg.	0.950	0.299	0.948	0.950	0.948	0.723	0.948	0.975	

=== Confusion Matrix ===

```

a  b  <-- classified as
194  3  |  a = Tepat
  8 16  |  b = Lambat
  
```

**Gambar 5.14** Klasifikasi *Naïve Bayes* (10 *Cross-Validation*)

Gambar 5.14 merupakan hasil klasifikasi *naïve bayes* pada *tools WEKA* dengan menggunakan 10 *Cross-Validation* yang menunjukkan hasil 210 prediksi benar dengan akurasi sebesar 95.0226% dan 11 prediksi salah dengan persentasi 4.9774% dengan waktu klasifikasi selama 0 detik.



#### 4. Hasil Klasifikasi *Naïve Bayes* Menggunakan *Tool WEKA* (70% *Percentage Split*)

Tes *Percentage Split* hasil klasifikasi akan dites dengan menggunakan k% dari data tersebut. Pada tes ini akan digunakan 70% *Percentage Split* dari data.

The screenshot displays the WEKA interface for a Naive Bayes classifier. The 'Test options' panel shows 'Percentage split' set to 70%. The 'Classifier output' panel provides a summary of performance metrics and a detailed accuracy breakdown by class.

**Summary Metrics:**

Metric	Value	Percentage
Correctly Classified Instances	61	92.4242 %
Incorrectly Classified Instances	5	7.5758 %
Kappa statistic	0.2534	
Mean absolute error	0.0744	
Root mean squared error	0.2203	
Relative absolute error	40.4833 %	
Root relative squared error	81.7289 %	
Total Number of Instances	66	

**Detailed Accuracy By Class:**

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.984	0.800	0.938	0.984	0.960	0.283	0.959	0.996	Tepat
	0.200	0.016	0.500	0.200	0.286	0.283	0.959	0.655	Lambat
Weighted Avg.	0.924	0.741	0.904	0.924	0.909	0.283	0.959	0.970	

**Confusion Matrix:**

```

=== Confusion Matrix ===
 a b  <-- classified as
60 1 | a = Tepat
 4 1 | b = Lambat

```

**Gambar 5.15** Klasifikasi *Naïve Bayes* (70% *Percentage Split*)

Gambar 5.15 merupakan hasil klasifikasi *naïve bayes* pada *tools WEKA* dengan menggunakan 70% *Percentage Split* yang menunjukkan hasil 61 prediksi benar dengan akurasi sebesar 92.4242% dan 5 prediksi salah dengan persentasi 7.5758% dengan waktu klasifikasi selama 0 detik.

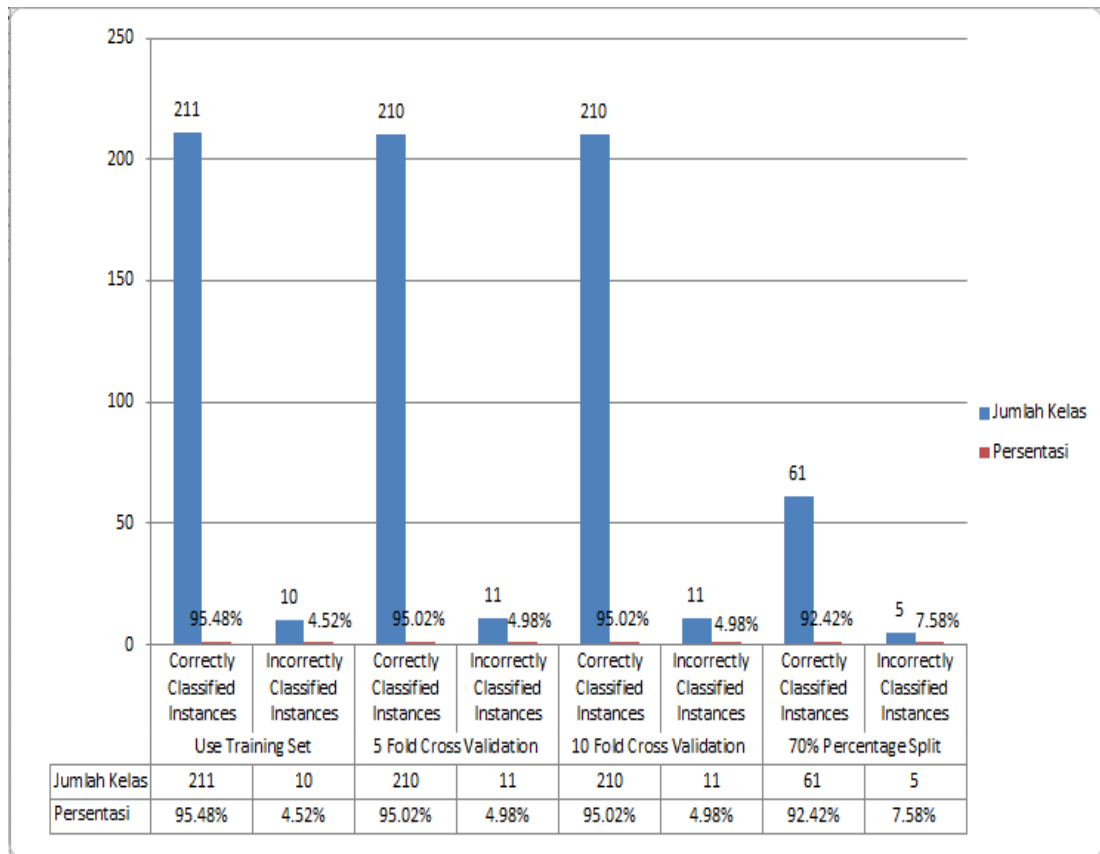
### 5.1.5 Hasil Perbandingan Evaluasi Akurasi *Naïve Bayes* Dari Data Kelulusan Mahasiswa Sistem Informasi

Setelah dilakukan analisis klasifikasi *naïve bayes* pada *tool WEKA* menggunakan *Use Training Set*, *5 Fold Cross Validation*, *10 Fold Cross Validation*, dan *70% Percentage Split*, maka diperoleh akurasi tertinggi yaitu dengan menggunakan *Use Training Set* dengan persentasi akurasi yaitu 95.4751% untuk *Correctly Classified Instances* dan 4.5249% untuk *Incorrectly Classified Instances*. Perbandingan hasil analisis dapat dilihat pada tabel 5.2.

**Tabel 5.2 Perbandingan Evaluasi Akurasi *Naïve Bayes***

Model Evaluasi	Akurasi	Jumlah Kelas	Persentasi
<i>Use Training Set</i>	<i>Correctly Classified Instances</i>	211	95.4751%
	<i>Incorrectly Classified Instances</i>	10	4.5249%
<i>5 Fold Cross Validation</i>	<i>Correctly Classified Instances</i>	210	95.0226%
	<i>Incorrectly Classified Instances</i>	11	4.9774%
<i>10 Fold Cross Validation</i>	<i>Correctly Classified Instances</i>	210	95.0226%
	<i>Incorrectly Classified Instances</i>	11	4.9774%
<i>70% Percentage Split</i>	<i>Correctly Classified Instances</i>	61	92.4242%
	<i>Incorrectly Classified Instances</i>	5	7.5758%

Grafik hasil akurasi klasifikasi *naïve bayes* menggunakan *tool WEKA* dapat dilihat pada gambar 5.16 :



**Gambar 5.16 Hasil Akurasi Klasifikasi Dengan *Naïve Bayes***

Gambar 5.16 merupakan grafik persentasi hasil klasifikasi *naïve bayes* dengan menggunakan 4 *test options*. Pada *Use Training Set* dengan jumlah kelas *Correctly Classified Instances* 211, *Incorrectly Classified Instances* 10, dan persentasi akurasi *Correctly Classified Instances* sebesar 95.4751%, *Incorrectly Classified Instances* 4.5249%. Pada tes *5 Fold Cross Validation* dengan jumlah kelas *Classified Instances* 210, *Incorrectly Classified Instances* 11, dan persentasi akurasi *Correctly Classified Instances* sebesar 95.0226%, *Incorrectly Classified Instances* 4.9774%. Pada tes *10 Fold Cross Validation* dengan jumlah kelas *Correctly Classified Instances* 210, *Incorrectly Classified Instances* 11, dan persentasi akurasi *Correctly Classified Instances* sebesar 95.0226%, *Incorrectly*

*Classified Instances* 4.9774%. Pada tes 70% *Percentage Split* dengan jumlah kelas *Correctly Classified Instances* 61, *Incorrectly Classified Instances* 5, dan persentasi akurasi *Correctly Classified Instances* sebesar 92.4242%, *Incorrectly Classified Instances* 7.5758%.

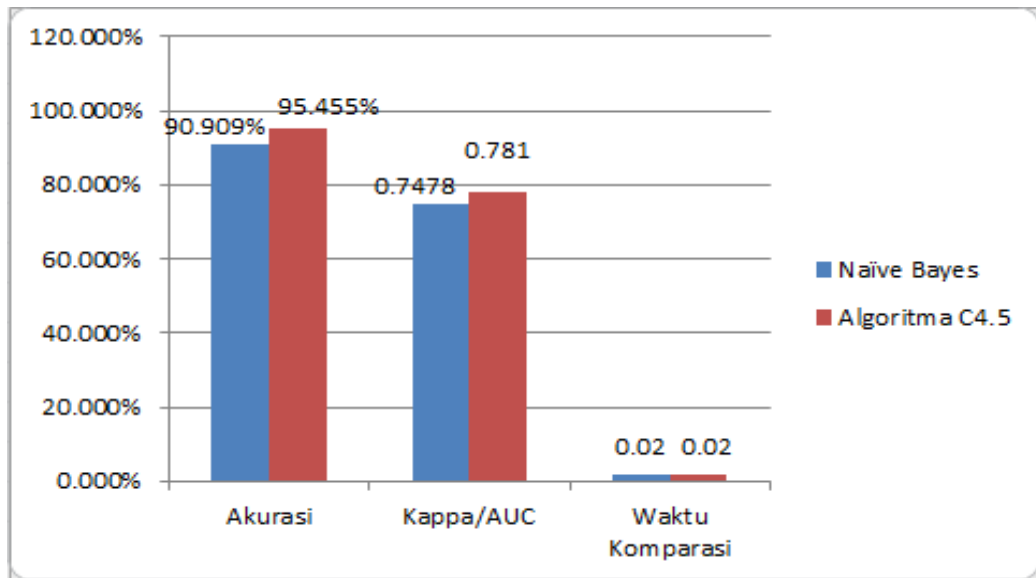
## 5.2 PERBANDINGAN HASIL KLASIFIKASI

Perbandingan hasil klasifikasi menggunakan 2 metode yaitu Algoritma C4.5 dan *Naïve Bayes* dapat dilihat pada tabel 5.3.

**Tabel 5.3 Perbandingan Hasil Klasifikasi**

Algoritma	<i>Naïve Bayes</i>	Algoritma C4.5
Akurasi	95.4751%	96.3801%
Kappa/AUC	0.7478	0.781
Waktu Komparasi	0.02 detik	0.02 detik

Pada tabel 5.3 diatas dapat diketahui metode yang memiliki nilai akurasi tinggi adalah metode Algoritma C4.5 dengan nilai akurasi 96.3801% dan AUC 0.781. Perbandingan Algoritma C4.5 dan *Naïve Bayes* juga dapat dilihat pada gambar 5.17.



Gambar 5.17 Perbandingan Hasil Klasifikasi

### 5.3 HASIL SELEKSI ATRIBUT MENGGUNAKAN CLASSIFIER ATTRIBUTE EVALUATION

The screenshot shows the WEKA software interface for attribute selection. The 'ClassifierAttributeEval' search method is selected, and the 'Ranker' search method is chosen. The 'Attribute Selection Mode' is set to 'Use full training set'. The 'Attribute selection output' window displays the following information:

```

Search Method:
  Attribute ranking.

Attribute Evaluator (supervised, Class (nominal): 7 Kelas Lulus):
  Classifier feature evaluator

Using Wrapper Subset Evaluator
Learning scheme: weka.classifiers.rules.ZeroR
Scheme options:
Subset evaluation: classification accuracy
Number of folds for accuracy estimation: 5

Ranked attributes:
0 6 Nilai D
0 5 Status Kuliah
0 2 Asal Sekolah
0 3 Jurusan Sekolah
0 4 IPK
0 1 JK

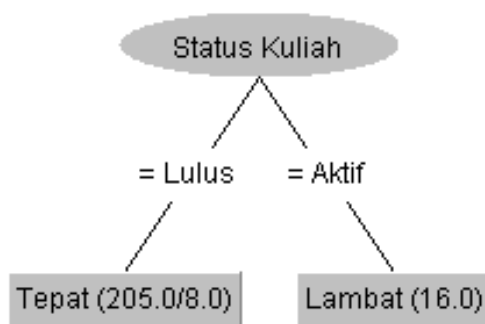
Selected attributes: 6,5,2,3,4,1 : 6
  
```

The 'Result list' shows the selected attributes: 6, 5, 2, 3, 4, 1.

Gambar 5.18 Hasil Seleksi Atribut di WEKA

Berdasarkan hasil seleksi atribut yang telah dilakukan menggunakan *Classifier Attribute Evaluation* maka urutan atribut yang paling mempengaruhi yaitu Nilai D, Status Kuliah, Asal Sekolah, Jurusan Sekolah, IPK, dan Jenis Kelamin.

#### 5.4 HASIL VISUALISASI *TREE* ALGORITMA C4.5 DENGAN WEKA



**Gambar 5.19** Visualisasi *Tree* Algoritma C4.5

Gambar 5.18 memiliki 2 jumlah daun dan 3 ukuran pohon. Node atau akar pertama adalah status kuliah dan mempunyai 2 (dua) variabel yaitu variabel lulus dan aktif. Pada kelas lulus mempunyai kelas tepat yang memiliki 205 data mahasiswa lulus dengan status kuliah tepat dan kelas lambat memiliki 16 data mahasiswa aktif dengan status kuliah lambat.