

BAB V

HASIL ANALISIS DAN VISUALISASI

5.1 PRAPROSES DATA

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

1. Data Mahasiswa Sistem Informasi

Data mentah ini merupakan semua data mahasiswa STIKOM Dinamika Bangsa Jambi jurusan SI sebelum dilakukannya penyeleksian data. Adapun data tersebut dapat dilihat pada tabel 5.1 :

Tabel 5.1 Data Mahasiswa Sistem Informasi

#	NIM	Nama	Status Kuliah	Mata Kuliah	SKS	Kontrak	NAS	Grade	Kredit	Periode
1	8040130001	Dewi Tursina	Lulus	Bahasa Inggris I	2	B	66	C+	5	2013
2	8040130001	Dewi Tursina	Lulus	Pengantar Manajemen	2	B	78	B+	7	2013
3	8040130001	Dewi Tursina	Lulus	Kalkulus	2	B	82	A	8	2013
4	8040130001	Dewi Tursina	Lulus	Pendidikan Agama	2	B	82	A	8	2013
5	8040130001	Dewi Tursina	Lulus	Pendidikan Pancasila	2	B	80	A	8	2013
6	8040130001	Dewi Tursina	Lulus	Pengantar Teknologi Informasi	3	B	71	B	9	2013
7	8040130001	Dewi Tursina	Lulus	Lab. Pengantar Tek. Informasi	1	B	80	A	4	2013
8	8040130001	Dewi Tursina	Lulus	Algoritma dan Pemrograman	3	B	70	B	9	2013
9	8040130001	Dewi Tursina	Lulus	Bahasa Indonesia	2	B	82	A	8	2013
10	8040130001	Dewi Tursina	Lulus	Lab. Algoritma dan Pemrograman	1	B	65	C+	2,5	2013
11	8040130001	Dewi Tursina	Lulus	Logika Matematika	2	B	62	C	4	2013
12	8040130001	Dewi Tursina	Lulus	Kecakapan Antar Personal	2	B	60	C	4	2013
13	8040130001	Dewi Tursina	Lulus	Etika Profesi	2	B	75	B+	7	2013
14	8040130001	Dewi Tursina	Lulus	Matematika Diskrit	2	B	70	B	6	2013
15	8040130001	Dewi Tursina	Lulus	Organisasi Komputer	2	B	68	C+	5	2013
16	8040130001	Dewi Tursina	Lulus	Sistem Operasi	2	B	55	D+	3	2013
17	8040130001	Dewi Tursina	Lulus	Pengantar Akuntansi	2	B	60	C	4	2013
18	8040130001	Dewi Tursina	Lulus	Sistem Digital	2	B	65	C+	5	2013
19	8040130001	Dewi Tursina	Lulus	Sistem Informasi	3	B	81	A	12	2013
20	8040130001	Dewi Tursina	Lulus	Struktur Data	2	B	70	B	6	2013
21	8040130001	Dewi Tursina	Lulus	Lab. Struktur Data	1	B	71	B	3	2013
22	8040130001	Dewi Tursina	Lulus	Pengetahuan Bisnis	2	B	80	A	8	2013
23	8040130001	Dewi Tursina	Lulus	Pendidikan Kewarganegaraan	2	B	74	B	6	2013
24	8040130001	Dewi Tursina	Lulus	Bahasa Inggris II	2	B	72	B	6	2013
25	8040130001	Dewi Tursina	Lulus	Lab. Pemrograman Berorientasi Objek I	1	B	60	C	2	2014

2. Seleksi Data Mahasiswa Sistem Informasi

Seleksi data mahasiswa sistem informasi ini merupakan proses penyeleksian data dengan memfokuskan pada atribut-atribut yang digunakan untuk perhitungan

seperti jenis kelamin, sekolah asal, jurusan sekolah asal, status kuliah, kriteria nilai, ipk dan kelas lulus. Adapun data tersebut dapat dilihat pada Tabel 5.2.

Tabel 5.2 Seleksi Data Mahasiswa Sistem Informasi

jenkel	sekolahasal	jurusan	statuskuliah	kriterianilai	ipk	kelaslulus
P	SMA	IPS	Lulus	Sangat Baik	Sedang	CEPAT
L	SMA	IPS	Lulus	Sangat Baik	Rendah	LAMBAT
P	SMA	IPS	Lulus	Sangat Baik	Sedang	TEPAT
P	SMK	Komputer	Lulus	Sangat Baik	Sedang	TEPAT
P	SMK	Ekonomi	Lulus	Sangat Baik	Tinggi	CEPAT
L	SMK	Komputer	Lulus	Sangat Baik	Sedang	TEPAT
P	SMK	Ekonomi	Lulus	Sangat Baik	Tinggi	CEPAT
P	SMK	Ekonomi	Lulus	Sangat Baik	Sedang	CEPAT
P	SMK	Perkantoran	Lulus	Sangat Baik	Sedang	CEPAT
L	SMK	Ekonomi	Lulus	Sangat Baik	Tinggi	CEPAT
L	SMK	IPS	Lulus	Sangat Baik	Tinggi	CEPAT
L	SMK	Ekonomi	Lulus	Sangat Baik	Sedang	CEPAT
P	SMK	Ekonomi	Lulus	Sangat Baik	Tinggi	CEPAT
P	SMK	Komputer	Lulus	Sangat Baik	Tinggi	CEPAT
P	SMK	Komputer	Lulus	Sangat Baik	Sedang	TEPAT
L	SMA	IPS	Lulus	Sangat Baik	Tinggi	CEPAT
P	SMK	Ekonomi	Lulus	Sangat Baik	Tinggi	CEPAT
P	SMA	IPS	Lulus	Sangat Baik	Tinggi	CEPAT
P	SMA	IPS	Lulus	Sangat Baik	Tinggi	CEPAT
P	SMA	IPS	Lulus	Sangat Baik	Sedang	CEPAT
L	SMK	Komputer	Lulus	Baik	Sedang	TEPAT
L	SMA	IPA	Lulus	Sangat Baik	Sedang	CEPAT
L	SMA	IPS	Lulus	Sangat Baik	Tinggi	CEPAT
P	SMA	IPS	Lulus	Sangat Baik	Tinggi	TEPAT
L	SMK	Komputer	Lulus	Sangat Baik	Sedang	TEPAT
L	SMA	IPS	Lulus	Cukup	Sedang	LAMBAT

3. Konversi Hasil Data

Data yang di di excel kemudian dipindahkan ke notepad kemudian di simpan dengan format .csv. ini dilakukan agar data dapat dimasukan kedalam *tools weka*, adapun bentuk data tersebut dapat dilihat pada gambar 5.1

```

jenkel,sekolahasal,jurusan,statuskuliah,kriterianilai,ipk,kelaslulus,
P ,SMA,IPS,Lulus ,Sangat Baik,Sedang,CEPAT,
L ,SMA,IPS,Lulus ,Sangat Baik,Rendah,LAMBAT,
P ,SMA,IPS,Lulus ,Sangat Baik,Sedang,TEPAT,
P ,SMK,Komputer,Lulus ,Sangat Baik,Sedang,TEPAT,
P ,SMK,Ekonomi,Lulus ,Sangat Baik,Tinggi,CEPAT,
L ,SMK,Komputer,Lulus ,Sangat Baik,Sedang,TEPAT,
P ,SMK,Ekonomi,Lulus ,Sangat Baik,Tinggi,CEPAT,
P ,SMK,Ekonomi,Lulus ,Sangat Baik,Sedang,CEPAT,
P ,SMK,Perkantoran,Lulus ,Sangat Baik,Sedang,CEPAT,
L ,SMK,Ekonomi,Lulus ,Sangat Baik,Tinggi,CEPAT,
L ,SMK,IPS,Lulus ,Sangat Baik,Tinggi,CEPAT,
L ,SMK,Ekonomi,Lulus ,Sangat Baik,Sedang,CEPAT,
P ,SMK,Ekonomi,Lulus ,Sangat Baik,Tinggi,CEPAT,
P ,SMK,Komputer,Lulus ,Sangat Baik,Tinggi,CEPAT,
P ,SMK,Komputer,Lulus ,Sangat Baik,Sedang,TEPAT,
L ,SMA,IPS,Lulus ,Sangat Baik,Tinggi,CEPAT,
P ,SMK,Ekonomi,Lulus ,Sangat Baik,Tinggi,CEPAT,
P ,SMA,IPS,Lulus ,Sangat Baik,Tinggi,CEPAT,
P ,SMA,IPS,Lulus ,Sangat Baik,Tinggi,CEPAT,
P ,SMA,IPS,Lulus ,Sangat Baik,Sedang,CEPAT,
L ,SMK,Komputer,Lulus ,Baik,Sedang,TEPAT,
L ,SMA,IPA,Lulus ,Sangat Baik,Sedang,CEPAT,
L ,SMA,IPS,Lulus ,Sangat Baik,Tinggi,CEPAT,
P ,SMA,IPS,Lulus ,Sangat Baik,Tinggi,TEPAT,
L ,SMK,Komputer,Lulus ,Sangat Baik,Sedang,TEPAT,
L ,SMA,IPS,Lulus ,Cukup,Sedang,LAMBAT,
L ,SMA,IPS,Lulus ,Sangat Baik,Rendah,CEPAT,
P ,SMA,IPS,Lulus ,Sangat Baik,Tinggi,CEPAT,
L ,SMK,Komputer,Lulus ,Baik,Sedang,CEPAT,
P ,SMK,Ekonomi,Lulus ,Sangat Baik,Sedang,CEPAT,
P ,SMK,Ekonomi,Lulus ,Sangat Baik,Sedang,CEPAT,
P ,SMK,Ekonomi,Lulus ,Sangat Baik,Tinggi,CEPAT,
P ,SMA,IPA,Lulus ,Sangat Baik,Sedang,CEPAT,
P ,SMA,IPA,Lulus ,Sangat Baik,Sedang,TEPAT,
P ,SMA,IPS,Lulus ,Sangat Baik,Sedang,CEPAT,
L ,SMK,Komputer,Lulus ,Baik,Sedang,TEPAT,

```

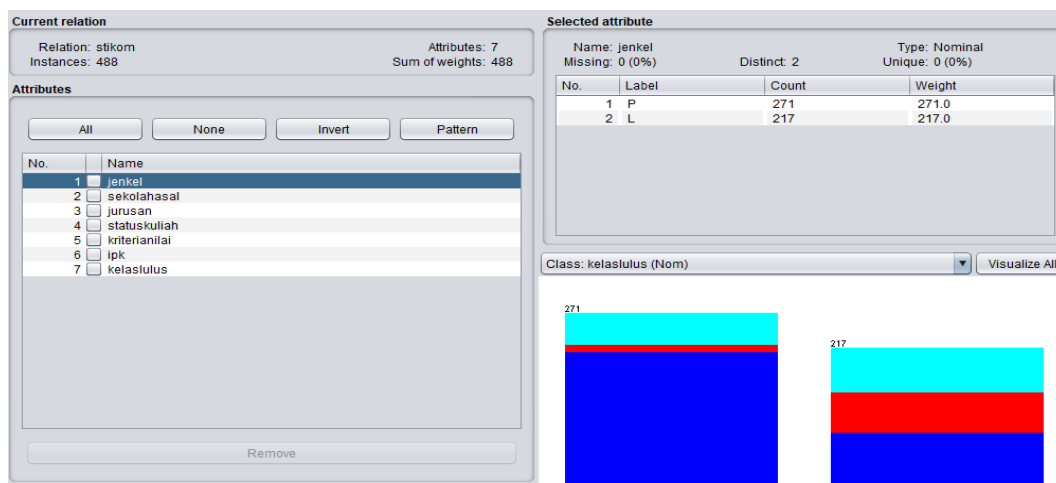
Gambar 5.1 Konversi Hasil Data

5.2 Hasil Visualisasi Data Setiap Atribut Dengan Menggunakan Weka

Tools weka digunakan untuk melihat hasil visualisasi setiap atribut dan digunakan untuk mengetahui hasil perhitungan menggunakan algoritma *C5.0*

1. Visualisasi Atribut Jenis Kelamin

Hasil visualisasi atribut tanggungan Jenis Kelamin dapat dilihat pada gambar 5.2

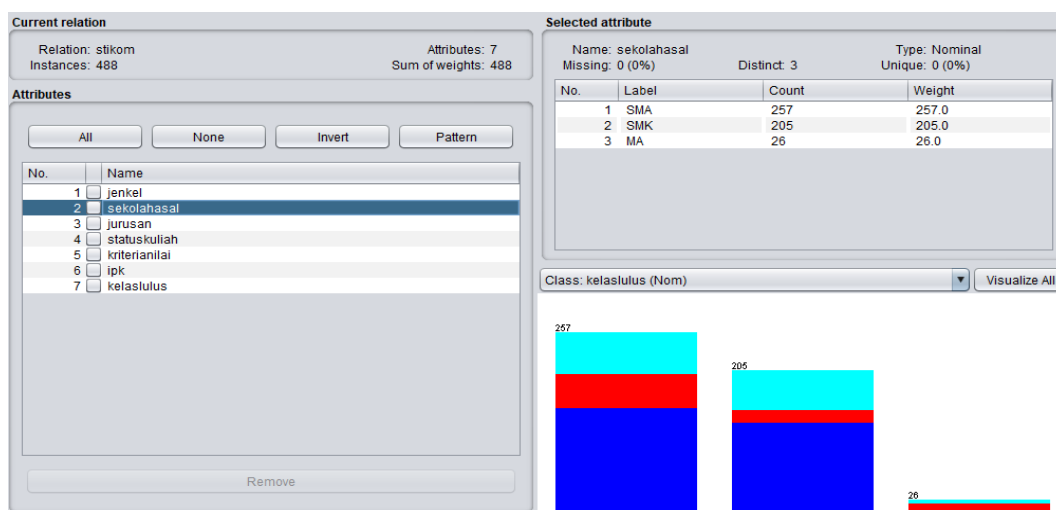


Gambar 5.2 Visualisasi Atribut Jenis Kelamin

Gambar 5.4 adalah visualisasi dari Atribut Jenis Kelamin. Sebanyak 488 data pada column selected atribut tidak terdapat *missing*. Distinct 10 Attribute unique 0 dengan jumlah mahasiswa Perempuan sebanyak 271 orang dan jumlah mahasiswa Laki-laki sebanyak 217 orang.

2. Visualisasi Atribut Sekolah Asal

Hasil visualisasi attribute Sekolah Asal dapat dilihat pada gambar 5.3

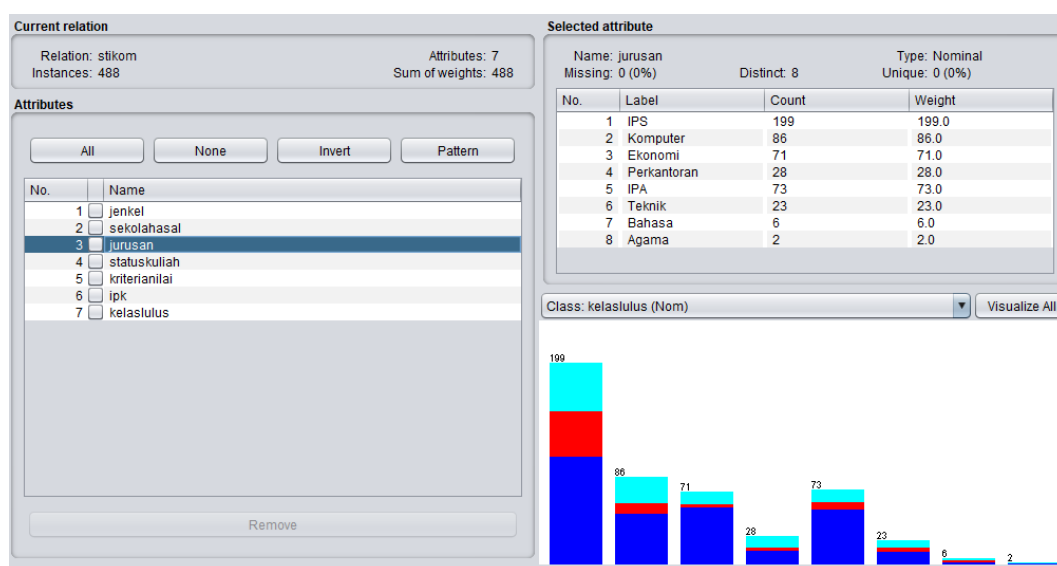


Gambar 5.3 Visualisasi Atribut Sekolah Asal

Gambar 5.3 adalah visualisasi dari Atribut Sekolah Asal. Sebanyak 488 data pada column selected atribut tidak terdapat *missing*. Distinct dengan nilai 3 Attribute unique 0, yang berasal dari SMA ada 257 orang, berasal dari SMK ada 205 orang dan yang berasal dar MA ada 26 orang.

3. Visualisasi Atribut Jurusan Sekolah Asal

Hasil visualisasi atribut lanjut usia dapat dilihat pada gambar 5.4

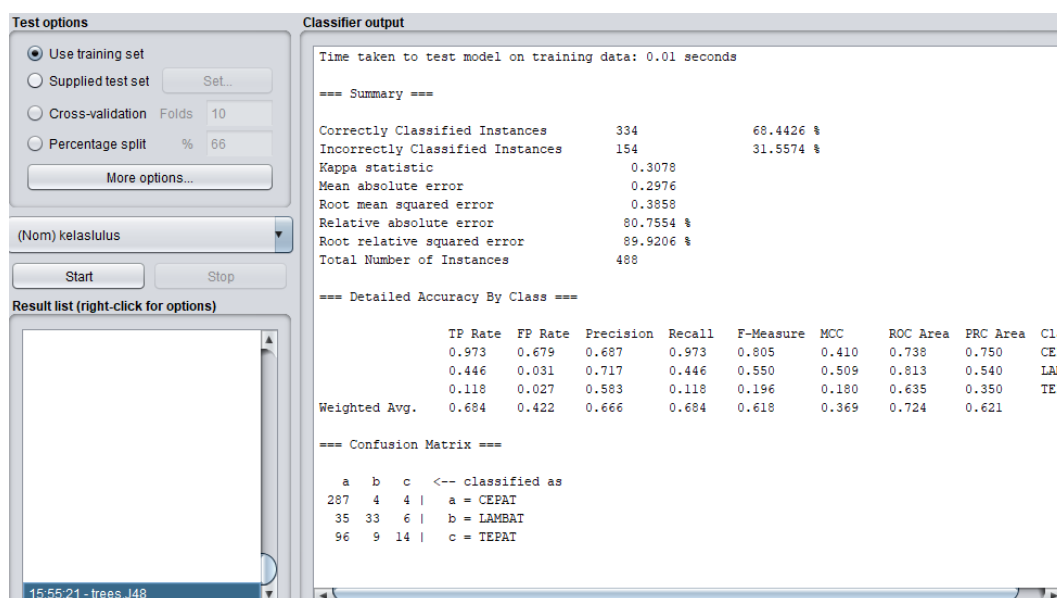


Gambar 5.4 Visualisasi Atribut Jurusan Sekolah Asal

Gambar 5.4 adalah visualisasi dari Atribut Jurusan Sekolah Asal. Sebanyak 488 data pada column selected atribut tidak terdapat *missing*. Distinct dengan nilai 8 Attribute unique 1, pada statistic minimum terdapat nilai 0, jurusan sekolah asal IPS berjumlah 199 orang, jurusan computer 85 orang, jurusan ekonomi 71 orang jurusan perkantoran 28 orang jurusan IPA 73 orang, jurusan Teknik 23 orang, jurusan Bahasa 6 orang dan jurusan agama 2 orang.

4. Hasil klasifikasi C5.0 menggunakan *tools weka (use data training)*

Test *Use Data Training* melakukan pengetesan data menggunakan data *training* itu sendiri. Dapat dilihat pada gambar 5.5



Gambar 5.5 Klasifikasi C5.0 (Use Data Training)

Gambar 5.5 merupakan hasil klasifikasi *c5.0* pada *tools WEKA* dengan menggunakan *use training set* yang menunjukkan hasil 334 prediksi benar dengan akurasi sebesar 68.4426% dan 154 prediksi salah dengan persentasi 31.5574% dengan waktu klasifikasi selama 0.01 detik.

5. Hasil klasifikasi C5.0 menggunakan *tools 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. Dapat dilihat pada gambar 5.6

The screenshot shows the WEKA Classifier window. The classifier selected is J48-C 0.25-M 2. Under 'Test options', 'Cross-validation' is selected with 'Folds' set to 5. The 'Classifier output' pane displays the following results:

```

=== Stratified cross-validation ===
=== Summary ===
Correctly Classified Instances      320           65.5738 %
Incorrectly Classified Instances    168           34.4262 %
Kappa statistic                     0.2517
Mean absolute error                  0.3101
Root mean squared error              0.4002
Relative absolute error              84.1083 %
Root relative squared error          93.28 %
Total Number of Instances           488

=== Detailed Accuracy By Class ===
          TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Cla
          0.969   0.679   0.686     0.969   0.803     0.403   0.727    0.748    CEP
          0.446   0.075   0.516     0.446   0.478     0.394   0.753    0.465    LAM
          0.008   0.016   0.143     0.008   0.016    -0.028  0.599    0.295    TEP
Weighted Avg.   0.656   0.426   0.528     0.656   0.562     0.297   0.700    0.595

=== Confusion Matrix ===
  a  b  c  <-- classified as
286  9  0 | a = CEPAT
 35 33  6 | b = LAMBAT
 96 22  1 | c = TEPAT

```

Gambar 5.6 Klasifikasi C5.0 (Use Data Training)

Gambar 5.6 merupakan hasil klasifikasi C5.0 pada tools WEKA dengan menggunakan *Cross Validation* yang menunjukkan hasil 320 prediksi benar dengan akurasi sebesar 65.5738% dan 168 prediksi salah dengan persentasi 34.4262% dengan waktu klasifikasi selama 0 detik.

6. Hasil klasifikasi C5.0 menggunakan tools weka (10 cross validation)

Test 10 *Cross-Validation* melakukan pengtesan 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 shows the WEKA Classifier interface. The classifier selected is J48 - C 0.25 - M 2. The test options are set to Cross-validation with 10 folds and a 66% percentage split. The classifier output shows the following results:

Time taken to build model: 0.05 seconds

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

Correctly Classified Instances	320	65.5738 %
Incorrectly Classified Instances	168	34.4262 %
Kappa statistic	0.2434	
Mean absolute error	0.3078	
Root mean squared error	0.4012	
Relative absolute error	83.4931 %	
Root relative squared error	93.5163 %	
Total Number of Instances	488	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.966	0.699	0.679	0.966	0.797	0.376	0.735	0.766	CEPAT
	0.392	0.048	0.592	0.392	0.472	0.410	0.737	0.454	LAMBAT
	0.050	0.035	0.316	0.050	0.087	0.034	0.590	0.299	TEPAT
Weighted Avg.	0.656	0.439	0.577	0.656	0.575	0.298	0.700	0.605	

=== Confusion Matrix ===

a	b	c	<-- classified as
285	6	4	a = CEPAT
36	29	9	b = LAMBAT

Gambar 5.7 Merupakan Hasil Klasifikasi C5.0 Pada Tools WEKA

Pada gambar 5.7 dengan menggunakan 10 *Cross-Validation* yang menunjukkan hasil 320 prediksi benar dengan akurasi sebesar 65.5738% dan 168 prediksi salah dengan persentasi 34.4262% dengan waktu klasifikasi selama 0.05 detik.

7. Hasil klasifikasi menggunakan *tools 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.

Classifier
Choose **J48 - C 0.25 - M 2**

Test options

Use training set
 Supplied test set Set...
 Cross-validation Folds 10
 Percentage split % 60
 More options...

(Nom) kelaslulus

Start Stop

Result list (right-click for options)

- 15:00:38 - trees.J48
- 15:26:37 - bayes.NaiveBayes
- 15:27:05 - trees.J48
- 15:55:21 - trees.J48
- 16:01:58 - trees.J48
- 16:23:03 - trees.J48
- 16:26:58 - trees.J48
- 16:30:00 - trees.J48
- 16:30:03 - trees.J48

Classifier output

Time taken to test model on test split: 0 seconds

=== Summary ===

Correctly Classified Instances	125	64.1026 %
Incorrectly Classified Instances	70	35.8974 %
Kappa statistic	0.2603	
Mean absolute error	0.3136	
Root mean squared error	0.3928	
Relative absolute error	84.0873 %	
Root relative squared error	89.1306 %	
Total Number of Instances	195	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.982	0.598	0.694	0.982	0.813	0.495	0.798	0.780	CEPAT
	0.184	0.006	0.875	0.184	0.304	0.355	0.826	0.572	LAMBAT
	0.159	0.132	0.259	0.159	0.197	0.032	0.620	0.279	TEPAT
Weighted Avg.	0.641	0.377	0.631	0.641	0.575	0.363	0.764	0.627	

=== Confusion Matrix ===

a	b	c	<-- classified as
111	0	2	a = CEPAT
13	7	18	b = LAMBAT
36	1	7	c = TEPAT

Gambar 5.8 Klasifikasi *C5.0* (60% *Percentage Split*)

Gambar 5.8 merupakan hasil klasifikasi *C5.0* pada *tools* WEKA dengan menggunakan 60% *Percentage Split* yang menunjukkan hasil 125 prediksi benar dengan akurasi sebesar 64.1026% dan 70 prediksi salah dengan persentasi 35.8974% dengan waktu klasifikasi selama 0 detik.

8. Hasil Klasifikasi *C5.0* Menggunakan *Tool WEKA* (80% *Percentage Split*)

Tes *Percentage Split* hasil klasifikasi akan dites dengan menggunakan k% dari data tersebut. Pada tes ini akan digunakan 80% *Percentage Split* dari data.

The screenshot shows the WEKA Classifier window. The classifier selected is J48-C 0.25-M 2. The test options are set to Percentage split at 80%. The classifier output shows a summary of performance metrics and a detailed accuracy by class table.

Classifier output Summary:

Metric	Value	Percentage
Correctly Classified Instances	61	62.2449 %
Incorrectly Classified Instances	37	37.7551 %
Kappa statistic	0.276	
Mean absolute error	0.3034	
Root mean squared error	0.3867	
Relative absolute error	79.3877 %	
Root relative squared error	86.0556 %	
Total Number of Instances	98	

Detailed Accuracy By Class:

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
CEPAT	0.981	0.523	0.697	0.981	0.815	0.547	0.823	0.781	CEPAT
LAMBAT	0.238	0.026	0.714	0.238	0.357	0.338	0.877	0.675	LAMBAT
TEPAT	0.130	0.160	0.200	0.130	0.158	-0.035	0.616	0.285	TEPAT
Weighted Avg.	0.622	0.331	0.584	0.622	0.563	0.366	0.786	0.642	

Confusion Matrix:

```

a b c <-- classified as
53 0 1 | a = CEPAT
 5 5 11 | b = LAMBAT
18 2 3 | c = TEPAT

```

Gambar 5.9 Klasifikasi *C5.0* (80% *Percentage Split*)

Gambar 5.9 merupakan hasil klasifikasi *C5.0* pada *tools* WEKA dengan menggunakan 80% *Percentage Split* yang menunjukkan hasil 61 prediksi benar dengan akurasi sebesar 62.2449% dan 37 prediksi salah dengan persentasi 37.7551% dengan waktu klasifikasi selama 0 detik.

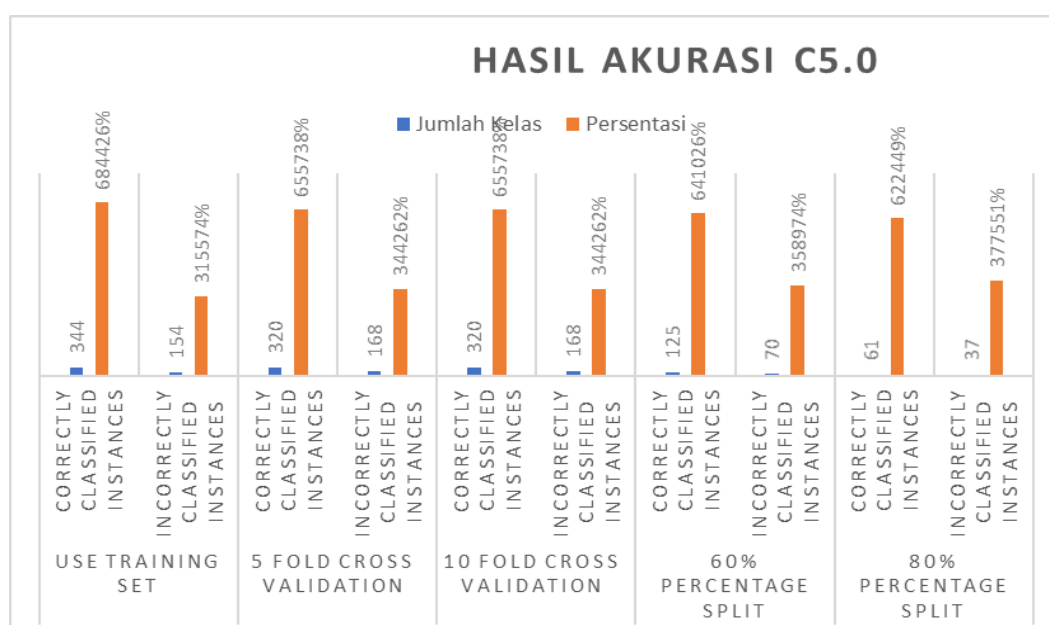
9. Hasil perbandingan evaluasi akurasi dari data mahasiswa stikom

Setelah dilakukan analisis klasifikasi *C5.0* 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 88.4426% untuk *Correctly Classified Instances* dan 31.5574% untuk *Incorrectly Classified Instances*. Perbandingan hasil analisis dapat dilihat pada tabel 5.3.

Tabel 5.3 Prbandingan Nilai Akurasi

Model Evaluasi	Akurasi	Jumlah Kelas	Persentasi
Use Training Set	<i>Correctly Classified Instances</i>	344	684426%
	<i>Incorrectly Classified Instances</i>	154	315574%
5 Fold Cross Validation	<i>Correctly Classified Instances</i>	320	655738%
	<i>Incorrectly Classified Instances</i>	168	344262%
10 Fold Cross Validation	<i>Correctly Classified Instances</i>	320	655738%
	<i>Incorrectly Classified Instances</i>	168	344262%
60% Percentage Split	<i>Correctly Classified Instances</i>	125	641026%
	<i>Incorrectly Classified Instances</i>	70	358974%
80% Percentage Split	<i>Correctly Classified Instances</i>	61	622449%
	<i>Incorrectly Classified Instances</i>	37	377551%

Grafik hasil akurasi klasifikasi C5.0 menggunakan tool WEKA dapat dilihat pada gambar 5.10

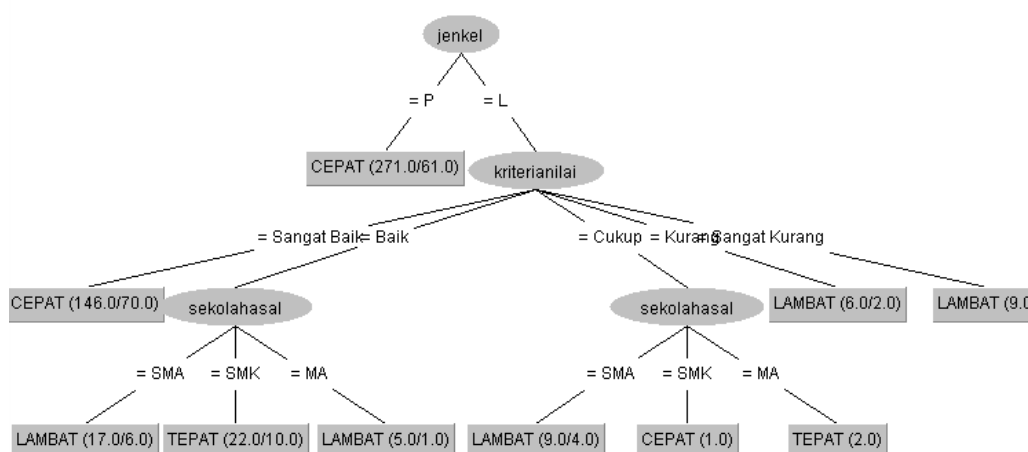


Gambar 5.10 Hasil Akurasi Metode C5.0

Gambar 5.10 merupakan grafik persentasi hasil klasifikasi *naive bayes* dengan menggunakan 5 *test options*. Pada *Use Training Set* dengan jumlah kelas *Correctly Classified Instances* 344, *Incorrectly Classified Instances* 154, dan persentasi akurasi *Correctly Classified Instances* sebesar 68.4426%, *Incorrectly Classified Instances* 31.5574%. Pada tes 5 *Fold Cross Validation* dengan jumlah kelas *Classified Instances* 320, *Incorrectly Classified Instances* 168, dan persentasi akurasi *Correctly Classified Instances* sebesar 65.5738%, *Incorrectly Classified Instances* 34.4262%. Pada tes 10 *Fold Cross Validation* dengan jumlah kelas *Correctly Classified Instances* 320, *Incorrectly Classified Instances* 168, dan persentasi akurasi *Correctly Classified Instances* sebesar 65.5738%, *Incorrectly Classified Instances* 34.4262%. Pada tes 60% *Percentage Split* dengan jumlah kelas *Correctly Classified Instances* 125, *Incorrectly Classified Instances* 70, dan persentasi akurasi *Correctly Classified Instances* sebesar 64.1026%, *Incorrectly Classified Instances* 35.8974%. Pada tes 80% *Percentage Split* dengan jumlah kelas *Correctly Classified Instances* 61, *Incorrectly Classified Instances* 37, dan persentasi akurasi *Correctly Classified Instances* sebesar 62.2449%, *Incorrectly Classified Instances* 37.7551%.

10. Pohon keputusan

Kemudian hasil klasifikasi tersebut digambarkan dalam bentuk pohon keputusan dapat dilihat pada gambar 5.11 Sebagai berikut:



Gambar 5.11 Centroid Cluster

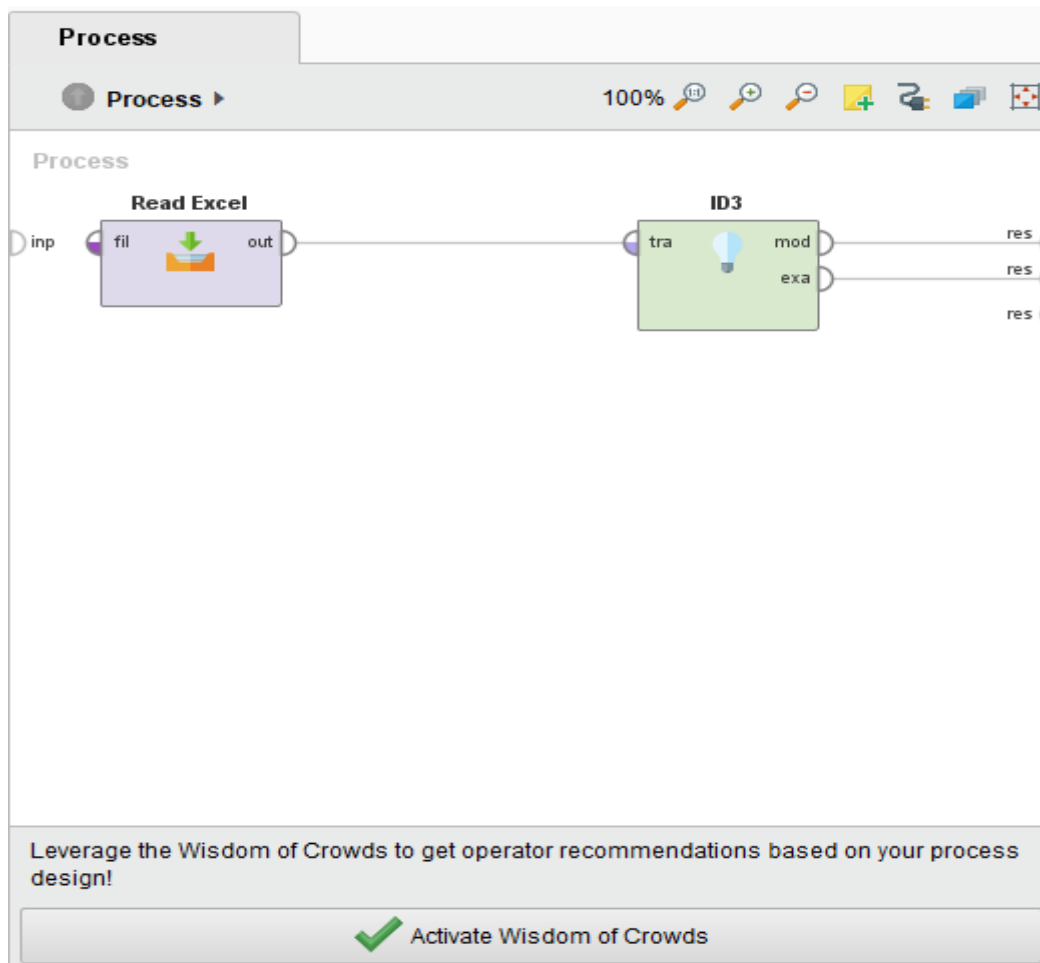
Dari gambar 5.11 dapat disimpulkan bahwa jenis kelamin perempuan yang lulus cepat sebanyak 271 orang. Untuk jenis kelamin laki-laki berkriteria sangat baik ada 146 orang, laki-laki berkriteria baik sekolah asal SMA dengan kelulusan lambat ada 17 orang, sekolah asal SMK kelulusan tepat ada 22 orang dan sekolah asal MA kelulusan lambat 5 orang. Laki-laki berkriteria cukup sekolah asalnya SMA dengan kelulusan lambat ada 9 orang, sekolah asal SMA kelulusan cepat 1 orang, sekolah asal orang MA kelulusan tepat 2 orang. Laki-laki berkriteria kurang dengan kelulusan lambat ada 6 orang, sedangkan laki-laki berkriteria sangat kurang berkelulusan lambat ada 9 orang.

5.3 Hasil Visualisasi Data Setiap Atribut Dengan Menggunakan Rapid Miner

Berikut merupakan bentuk visualisasi hasil dengan menggunakan tools *Rapid Miner*, yaitu :

1. Visualisasi proses metode *C5.0*

Visualisasi proses metode *C5.0* pada tools *Rapid Miner* dapat dilihat pada tabel 5.12



Gambar 5.12 Visualisasi Proses Metode *C5.0*

Pada gambar 5.12 dijelaskan bagaimana cara proses *ID3* klasifikasi *C5.0* menggunakan RapidMiner, langkah-langkah tersebut diantaranya :

- a. Pilih *import* data untuk memilih data mana yang akan di proses.
- b. Kemudian pilih *My Computer*, dan cari data yang akan di proses.
- c. Kemudian lanjutkan kelangkah berikutnya *next* sampai *finish* agar data tersebut bisa di olah.

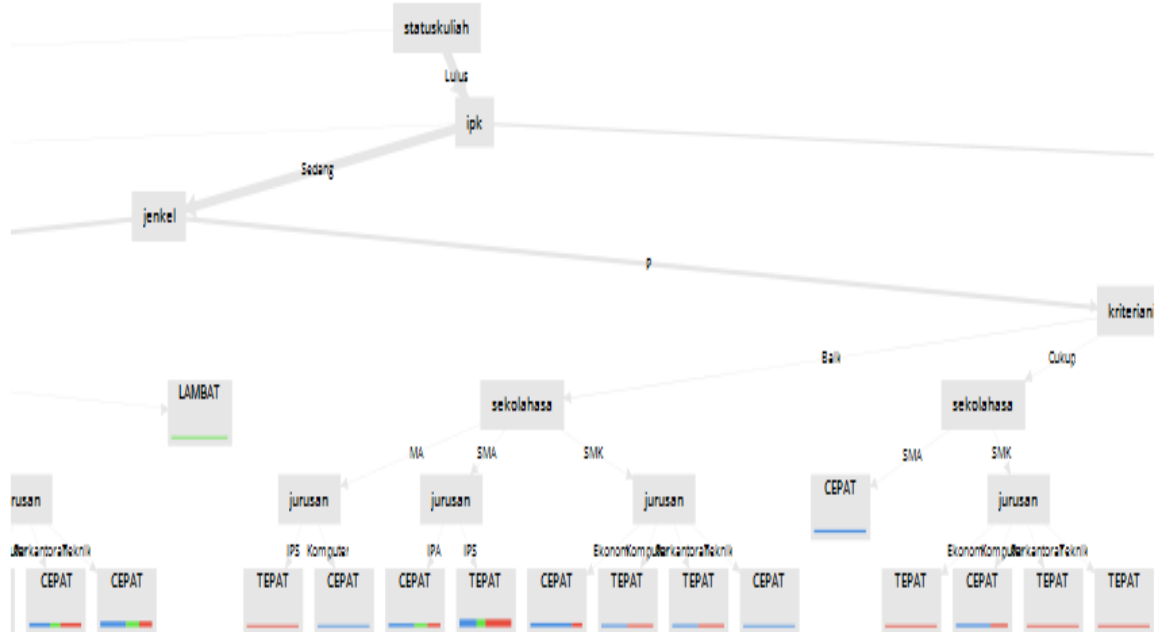
- d. Pada *coloumn Repository* ada beberapa *list*, kemudian pilih list *Local Repository*, didalam *list Repository* terdapat data yang telah di *import* dari *excel* ke *RapidMiner*, kemudian arahkan kursor tersebut ke data yang akan di olah, tahan dan tarik data tersebut ke arah *coloumn process*.
 - e. Pada *coloumn Operators* terdapat *search box* setelah itu klik dan cari metode apa yang akan kita gunakan, yaitu metode *ID3*, kemudian arahkan kursor tersebut ke metode *ID3*, tahan dan tarik metode *ID3* tersebut ke arah *coloumn process*.
 - f. Pada *coloumn process* terdapat 2 proses yang akan dilakukan perhitungannya, diantaranya *read excel*, dan *ID3*. Hubungkan 2 proses tersebut dengan cara, pada *read excel* terdapat (out) *output* yang kemudian di hubungkan ke (exa) *example set* yang tertera pada proses *clustering*. Kemudian pada tabel proses terdapat 2 fungsi proses, yang pertama *read excel* hubungkan ke *ID3* kemudian hubungkan ke res agar pencarian dapat dilanjutkan.
 - g. Setelah semua fungsi terhubung, lakukan run untuk memulai proses *clasifikasi*.
2. Visualisasi data klasifikasi
- Visualisasi klasifikasi pada *Rapid Miner* dapat dilihat pada gambar 5.13

Open in Turbo Prep Auto Model Filter (488 /

Row No.	kelaslulus	jenkel	sekolahasal	jurusan	statuskuliah	kriterianilai	ipk
1	CEPAT	P	SMA	IPS	Lulus	Sangat Baik	Sedang
2	LAMBAT	L	SMA	IPS	Lulus	Sangat Baik	Rendah
3	TEPAT	P	SMA	IPS	Lulus	Sangat Baik	Sedang
4	TEPAT	P	SMK	Komputer	Lulus	Sangat Baik	Sedang
5	CEPAT	P	SMK	Ekonomi	Lulus	Sangat Baik	Tinggi
6	TEPAT	L	SMK	Komputer	Lulus	Sangat Baik	Sedang
7	CEPAT	P	SMK	Ekonomi	Lulus	Sangat Baik	Tinggi
8	CEPAT	P	SMK	Ekonomi	Lulus	Sangat Baik	Sedang
9	CEPAT	P	SMK	Perkantoran	Lulus	Sangat Baik	Sedang
10	CEPAT	L	SMK	Ekonomi	Lulus	Sangat Baik	Tinggi
11	CEPAT	L	SMK	IPS	Lulus	Sangat Baik	Tinggi
12	CEPAT	L	SMK	Ekonomi	Lulus	Sangat Baik	Sedang
13	CEPAT	P	SMK	Ekonomi	Lulus	Sangat Baik	Tinggi
14	CEPAT	P	SMK	Komputer	Lulus	Sangat Baik	Tinggi
15	TEPAT	P	SMK	Komputer	Lulus	Sangat Baik	Sedang

Gambar 5.13 Visualisasi Letak Data dan Klasifikasi

Pada gambar 5.13 menjelaskan data dari klasifikasi menggunakan metode c5.0 yang akan dirubah dalam bentuk pohon keputusan seperti gambar 5.14



Gambar 5.14 Visualisasi Pohon Keputusan

Pada gambar 5.14 menampilkan pohon keputusan hasil klasifikasi yang akan di jelaskan pad gambar 5.15

Tree

```

statuskuliah = Aktif : LAMBAT {CEPAT=0, LAMBAT=15, TEPAT=0}
statuskuliah = Lulus
| ipk = Rendah
| | kriterianilai = Baik: LAMBAT {CEPAT=0, LAMBAT=2, TEPAT=0}
| | kriterianilai = Cukup: LAMBAT {CEPAT=0, LAMBAT=1, TEPAT=0}
| | kriterianilai = Sangat Baik
| | | jenkel = L
| | | | sekolahasal = SMA
| | | | | jurusan = IPS: LAMBAT {CEPAT=1, LAMBAT=1, TEPAT=0}
| ipk = Sedang
| | jenkel = L
| | | kriterianilai = Baik
| | | | sekolahasal = MA
| | | | | jurusan = IPA: TEPAT {CEPAT=0, LAMBAT=0, TEPAT=1}
| | | | | jurusan = IPS: LAMBAT {CEPAT=0, LAMBAT=2, TEPAT=0}
| | | | | jurusan = Teknik: LAMBAT {CEPAT=0, LAMBAT=1, TEPAT=0}
| | | | | sekolahasal = SMA
| | | | | | jurusan = IPA: LAMBAT {CEPAT=1, LAMBAT=1, TEPAT=0}
| | | | | | jurusan = IPS: LAMBAT {CEPAT=2, LAMBAT=7, TEPAT=3}
| | | | | | jurusan = Komputer: LAMBAT {CEPAT=0, LAMBAT=1, TEPAT=0}
| | | | | sekolahasal = SMK
| | | | | | jurusan = Ekonomi: TEPAT {CEPAT=1, LAMBAT=0, TEPAT=2}
| | | | | | jurusan = Komputer: TEPAT {CEPAT=2, LAMBAT=4, TEPAT=5}
| | | | | | jurusan = Perkantoran: TEPAT {CEPAT=0, LAMBAT=0, TEPAT=2}
| | | | | | | jurusan = Teknik: TEPAT {CEPAT=1, LAMBAT=1, TEPAT=3}
| | | | kriterianilai = Cukup
| | | | | sekolahasal = MA: TEPAT {CEPAT=0, LAMBAT=0, TEPAT=2}
| | | | | sekolahasal = SMA

```

```

| | | | | jurusan = Bahasa: LAMBAT {CEPAT=0, LAMBAT=1, TEPAT=0}
| | | | | jurusan = IPS: LAMBAT {CEPAT=1, LAMBAT=3, TEPAT=2}
| | | | | jurusan = Komputer: TEPAT {CEPAT=0, LAMBAT=0, TEPAT=1}
| | | | | sekolahhasal = SMK: CEPAT {CEPAT=1, LAMBAT=0, TEPAT=0}
| | | | | kriterianilai = Kurang
| | | | | sekolahhasal = MA: LAMBAT {CEPAT=0, LAMBAT=1, TEPAT=0}
| | | | | sekolahhasal = SMA
| | | | | | jurusan = IPS: TEPAT {CEPAT=0, LAMBAT=1, TEPAT=2}
| | | | | kriterianilai = Sangat Baik
| | | | | sekolahhasal = MA
| | | | | | jurusan = Agama: CEPAT {CEPAT=1, LAMBAT=0, TEPAT=0}
| | | | | | jurusan = IPA: TEPAT {CEPAT=0, LAMBAT=0, TEPAT=1}
| | | | | | jurusan = IPS: CEPAT {CEPAT=2, LAMBAT=0, TEPAT=0}
| | | | | sekolahhasal = SMA
| | | | | | jurusan = IPA: TEPAT {CEPAT=4, LAMBAT=1, TEPAT=5}
| | | | | | jurusan = IPS: TEPAT {CEPAT=21, LAMBAT=12, TEPAT=22}
| | | | | sekolahhasal = SMK
| | | | | | jurusan = Ekonomi: CEPAT {CEPAT=7, LAMBAT=3, TEPAT=4}
| | | | | | jurusan = Komputer: CEPAT {CEPAT=15, LAMBAT=2, TEPAT=10}
| | | | | | jurusan = Perkantoran: CEPAT {CEPAT=2, LAMBAT=1, TEPAT=2}
| | | | | | jurusan = Teknik: CEPAT {CEPAT=4, LAMBAT=2, TEPAT=2}
| | | | | kriterianilai = Sangat Kurang: LAMBAT {CEPAT=0, LAMBAT=1, TEPAT=0}
| | | | | jenkel = P
| | | | | kriterianilai = Baik
| | | | | sekolahhasal = MA
| | | | | | jurusan = IPS: TEPAT {CEPAT=0, LAMBAT=0, TEPAT=1}
| | | | | | jurusan = Komputer: CEPAT {CEPAT=1, LAMBAT=0, TEPAT=0}
| | | | | sekolahhasal = SMA
| | | | | | jurusan = IPA: CEPAT {CEPAT=2, LAMBAT=1, TEPAT=1}
| | | | | | jurusan = IPS: TEPAT {CEPAT=4, LAMBAT=2, TEPAT=6}
| | | | | sekolahhasal = SMK
| | | | | | jurusan = Ekonomi: CEPAT {CEPAT=4, LAMBAT=0, TEPAT=1}

```

```

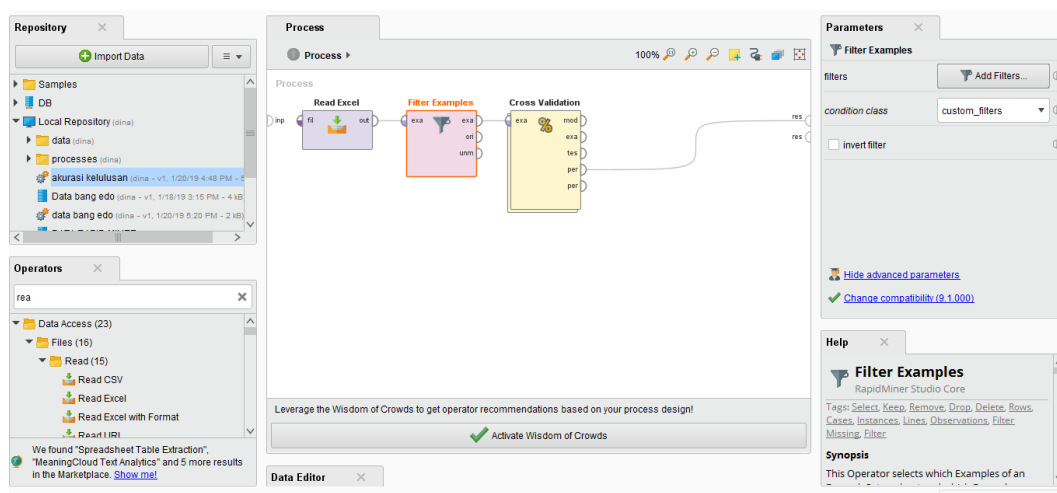
| | | | sekolahhasal = SMK
| | | | | jurusan = Ekonomi: CEPAT {CEPAT=4, LAMBAT=0, TEPAT=1}
| | | | | jurusan = Komputer: TEPAT {CEPAT=1, LAMBAT=0, TEPAT=1}
| | | | | jurusan = Perkantoran: TEPAT {CEPAT=1, LAMBAT=0, TEPAT=1}
| | | | | jurusan = Teknik: CEPAT {CEPAT=1, LAMBAT=0, TEPAT=0}
| | | kriterianilai = Cukup
| | | | sekolahhasal = SMA: CEPAT {CEPAT=3, LAMBAT=0, TEPAT=0}
| | | | sekolahhasal = SMK
| | | | | jurusan = Ekonomi: TEPAT {CEPAT=0, LAMBAT=0, TEPAT=1}
| | | | | jurusan = Komputer: CEPAT {CEPAT=2, LAMBAT=0, TEPAT=1}
| | | | | jurusan = Perkantoran: TEPAT {CEPAT=0, LAMBAT=0, TEPAT=1}
| | | | | jurusan = Teknik: TEPAT {CEPAT=0, LAMBAT=0, TEPAT=1}
| | | kriterianilai = Sangat Baik
| | | | jurusan = Agama: TEPAT {CEPAT=0, LAMBAT=0, TEPAT=1}
| | | | jurusan = Bahasa
| | | | | sekolahhasal = MA: CEPAT {CEPAT=1, LAMBAT=0, TEPAT=0}
| | | | | sekolahhasal = SMA: TEPAT {CEPAT=0, LAMBAT=0, TEPAT=2}
| | | | jurusan = Ekonomi
| | | | | sekolahhasal = SMK: CEPAT {CEPAT=16, LAMBAT=0, TEPAT=4}
| | | | jurusan = IPA
| | | | | sekolahhasal = MA: CEPAT {CEPAT=2, LAMBAT=0, TEPAT=0}
| | | | | sekolahhasal = SMA: CEPAT {CEPAT=20, LAMBAT=2, TEPAT=3}
| | | | jurusan = IPS
| | | | | sekolahhasal = MA: CEPAT {CEPAT=2, LAMBAT=1, TEPAT=0}
| | | | | sekolahhasal = SMA: CEPAT {CEPAT=34, LAMBAT=2, TEPAT=9}
| | | | jurusan = Komputer
| | | | | sekolahhasal = SMK: CEPAT {CEPAT=17, LAMBAT=0, TEPAT=6}
| | | | jurusan = Perkantoran
| | | | | sekolahhasal = SMK: CEPAT {CEPAT=9, LAMBAT=1, TEPAT=5}
| | | | jurusan = Teknik
| | | | | sekolahhasal = SMK: CEPAT {CEPAT=5, LAMBAT=0, TEPAT=1}
| ipk = Tinggi
| | jenkel = L
| | | sekolahhasal = SMA: CEPAT {CEPAT=13, LAMBAT=0, TEPAT=0}
| | | sekolahhasal = SMK
| | | | jurusan = Ekonomi: CEPAT {CEPAT=3, LAMBAT=0, TEPAT=0}
| | | | jurusan = IPS: CEPAT {CEPAT=1, LAMBAT=0, TEPAT=0}
| | | | jurusan = Komputer
| | | | | kriterianilai = Sangat Baik: LAMBAT {CEPAT=1, LAMBAT=1, TEPAT=1}
| | | | jurusan = Teknik: CEPAT {CEPAT=1, LAMBAT=0, TEPAT=0}
| | jenkel = P
| | | jurusan = Ekonomi: CEPAT {CEPAT=25, LAMBAT=0, TEPAT=0}
| | | jurusan = IPA
| | | | sekolahhasal = SMA
| | | | | kriterianilai = Sangat Baik: CEPAT {CEPAT=20, LAMBAT=0, TEPAT=1}
| | | jurusan = IPS
| | | | sekolahhasal = MA: CEPAT {CEPAT=2, LAMBAT=0, TEPAT=0}
| | | | sekolahhasal = SMA
| | | | | kriterianilai = Sangat Baik: CEPAT {CEPAT=26, LAMBAT=0, TEPAT=1}
| | | jurusan = Komputer
| | | | sekolahhasal = SMK
| | | | | kriterianilai = Sangat Baik: CEPAT {CEPAT=10, LAMBAT=0, TEPAT=1}
| | | jurusan = Perkantoran: CEPAT {CEPAT=2, LAMBAT=0, TEPAT=0}

```

Gambar 5.15 Visualisasi Penjelasan Pohon Keputusan

3. Visualisasi nilai akurasi

Adapun visualisasi nilai akurasi metode *C5.0* dapat dilihat pada gambar 5.16



Gambar 5.16 Visualisasi Akurasi Metode *C5.0*

Pada gambar 5.16 dijelaskan bagaimana cara proses mendapatkan nilai akurasi metode *C5.0* menggunakan RapidMiner, langkah-langkah tersebut diantaranya :

- Cari read by excel pada kolom perator kemudian tarik ke kolom process dan masukan data yang ingin di proses.
- Kemudian cari filter example dan tarik ke kolom proses
- Kemudian tambahkan operator cross validation ke kolom proses
- Koneksikan setiap operator ke operator lain pada cross validation ambil bagian per dan hubungkan ke res
- Setelah semuanya terhubung lakukan proses run, maka akan muncul tampilan seperti gambar 5.17

Result History PerformanceVector (Performance (2))

Criterion: accuracy, kappa

Table View Plot View

accuracy: 64.97% +/- 5.94% (micro average: 64.96%)

	true CEPAT	true LAMBAT	true TEPAT	class precision
pred. CEPAT	254	26	78	70.95%
pred. LAMBAT	8	33	11	63.46%
pred. TEPAT	33	15	30	38.46%
class recall	86.10%	44.59%	25.21%	

Gambar 5.17 Visualisasi Akurasi Metode C5.0

Pada gambar 5.17 menjelaskan bahwa akurasi perhitungan c5.0 adalah sebesar 64.97%.

5.4 PERBANDINGAN HASIL KLASIFIKASI NAIVE BAYES

Perbandingan hasil klasifikasi *naive bayes* menggunakan 3 tool yaitu *Excel*, *WEKA*, dan *RapidMiner* dapat dilihat pada tabel 5.4.

Tabel 5.4 Perbandingan Hasil Klasifikasi C5.0

Perbandingan		<i>Excel</i>	<i>WEKA</i>	<i>RapidMiner</i>
Jumlah	Kelas “CEPAT”	362	287	254
Jumlah	Kelas “TEPAT”	33	33	33
Jumlah	Kelas “LAMBAT”	48	14	30

Perbandingan	<i>Excel</i>	<i>WEKA</i>	<i>RapidMiner</i>
Akurasi	64,139%	68.4426% (<i>Correctly Classified Instance</i>) dan 31.5574% (<i>Incorrectly Classified Instance</i>)	64.97%

Pada tabel 5.4 diperoleh perbandingan hasil perhitungan klasifikasi prediksi lama masa studi mahasiswa menggunakan algoritma *C5.0* dengan menggunakan keseluruhan *data training* menjadi *data testing* pada *tool Excel*, *WEKA*, dan *RapidMiner*. Dapat disimpulkan bahwa perhitungan menggunakan *tool Excel* dan *RapidMiner* memiliki hasil akurasi yang sama, pada *tool Excel* kelas “CEPAT” berjumlah 362 mahasiswa, pada kelas “TEPAT” berjumlah 33 mahasiswa dan pada kelas “LAMBAT” berjumlah 48 mahasiswa dan pada *tool RapidMiner* kelas “CEPAT” berjumlah 254 mahasiswa, pada kelas “TEPAT” berjumlah 33 mahasiswa dan pada kelas “LAMBAT” berjumlah 30 mahasiswa, sedangkan pada *tool WEKA* kelas “CEPAT” berjumlah 287 mahasiswa, pada kelas “TEPAT” berjumlah 33, dan pada kelas “LAMAT” berjumlah 14 mahasiswa. Hasil akurasi pada *tool Excel* yaitu sebesar 64,139%, sedangkan pada *tool WEKA* *Correctly Classified Instance* sebesar 68.4426% dan *Incorrectly Classified Instance* sebesar 31.5574% dan pada *Rapid Miner* sebesar 64,97%.