

## PENERAPAN ALGORITMA C4.5 UNTUK PENENTUAN KELAYAKAN PEMBERIAN KREDIT (Studi Kasus : Koperia - Koperasi Warga Komplek Gandaria)

### *Implementation of C4.5 Algorithm To Determine The Feasibility of Loan (Case Study: Koperia – Koperasi Warga Komplek Gandaria)*

Teguh Budi Santoso, [teguh.santos12@gmail.com](mailto:teguh.santos12@gmail.com)<sup>1)</sup>, Dela Sekardiana,  
[dellasekar23@gmail.com](mailto:dellasekar23@gmail.com)<sup>2)</sup>

<sup>1)2)</sup> <sup>1)</sup>Teknik Informatika/Fakultas Teknik, Universitas Satya Negara Indonesia

#### ABSTRACT

*Current credit giving in KOPERIA (Koperasi Warga Komplek Gandaria) is still based on an objective process. Difficulties in determining the feasibility of giving credit are often experienced by cooperative managers, so that problems arise in the cooperative is a default payment of credit installments of customers in KOPERIA. This study aims to form a decision tree classification model to determine the customer's credit worthiness. In this study the application of C4.5 Algorithm, based on the Sets and Attributes used in this study, namely, the amount of income divided into 2 categories > 5 million and 3-5 million, the amount of balance divided into three, namely > 3 million, 1-3 million and <1 Million, The Loan Amount is divided into three, namely 1-4 Months, 5-8 months, and 9-12 Months and Requirements with attributes of Business Capital, buying goods and others. In this study determine the appropriate root nodes, the classification results using C4.5 Algorithm shows that the accuracy of 97.5% is obtained, based on the results obtained shows that the c4.5 algorithm is suitable to be used to determine the feasibility of lending customers to KOPERIA.*

**Keywords:** Data Mining, C4.5 Algorithm, loan feasibility

#### ABSTRAK

Pemberian kredit saat ini yang terjadi di KOPERIA (Koperasi Warga Komplek Gandaria) kepada nasabah masih berdasarkan proses yang tidak obyektif. Kesulitan dalam menentukan kelayakan pemberian kredit yang sering dialami oleh pengurus koperasi, sehingga muncul masalah pada koperasi adalah macetnya pembayaran angsuran kredit nasabah pada KOPERIA. Penelitian ini bertujuan membentuk model klasifikasi pohon keputusan untuk menentukan kelayakan kredit nasabah. Pada penelitian ini penerapan Algoritma C4.5, berdasarkan Himpunan dan Atribut yang digunakan pada penelitian ini yaitu, Jumlah Penghasilan dibagi menjadi 2 kategori >5 Juta dan 3-5Juta, Jumlah Saldo dibagi menjadi tiga yaitu >3 Juta, 1-3 Juta dan < 1Juta, Jumlah Pinjaman dibagi menjadi tiga yaitu 1-4 Bulan, 5-8 bulan, dan 9 -12 Bulan dan Keperluan dengan attribute Modal Usaha, membeli barang dan lain-lain. Pada Penelitian ini menentukan node akar yang sesuai, hasil klasifikasi menggunakan Algoritma C4.5 menunjukkan bahwa diperoleh akurasi 97,5%, berdasarkan hasil yang diperoleh menunjukkan bahwa algoritma c4.5 cocok digunakan untuk menentukan kelayakan pemberian kredit nasabah pada KOPERIA.

**Kata Kunci:** Data Mining, Algoritma C4.5, kelayakan kredit

#### PENDAHULUAN

Koperia adalah sebuah badan usaha yang bergerak dibidang simpan pinjam uang untuk memenuhi kebutuhan masyarakat dalam hal keuangan. Dapat diketahui bahwa di koperia sudah banyak

yang melakukan transaksi kredit akan tetapi pengurus koperia pada saat ini untuk memberikan kredit kepada nasabah masih berdasarkan proses yang tidak obyektif. Oleh sebab itu sulit menentukan kelayakan pemberian kredit yang sering dialami oleh pengurus koperia. Adapun masalah yang

muncul pada koperia adalah pembayaran angsuran. Seseorang yang sering menunggak dikarenakan penghasilan mereka yang kurang mencukupi. Dimana jika banyak nasabah yang menunggak dalam pembayaran maka akan mengganggu sistem keuangan perusahaan.

Untuk penentuan kelayakan pemberian kredit bisa dilakukan dengan proses perhitungan yang tepat (Firmansyah,2011), dengan menggunakan algoritma C4.5, dapat menentukan model pohon keputusan (Teguh Budi Santoso,2014) untuk meningkatkan akurasi dalam menganalisa kelayakan kredit yang diajukan calon debitur, Dalam metode ini diharapkan akan mampu untuk menentukan pemberian kredit yang layak atau tidak layak kepada nasabah. Dengan menerapkan beberapa atribut yang mampu mengurangi tingkat resiko penunggakan. Adapun atribut yang digunakan yaitu jumlah penghasilan, jumlah saldo, jumlah pinjaman, jangka waktu dan keperluan. Penelitian ini bertujuan untuk penentuan kelayakan pemberian kredit kepada nasabah, sehingga pihak koperasi dapat membantu menyelesaikan penentuan kelayakan pemberian kredit.

Permasalahan yang muncul berdasar kan uraian diatas adalah masih terjadinya proses yang sifatnya objektif dalam penentuan kelayakan kredit, yang menyebabkan tidak akuratnya data, seperti pengajuan kredit yang bermasalah dalam pembayaran angsuran kredit.

Dengan menerapkan algoritma C4.5 diharapkan dapat meningkatkan keakuratan analisa kredit, dalam penentuan kelayakan kredit.

## METODOLOGI PENELITIAN

Pada penelitian ini data yang digunakan adalah data sekunder atau data yang berasal dari jurnal, paper, buku serta beberapa informasi lainnya yang berhubungan dengan penelitian yang dikutip.

Pengolahan data yang dibutuhkan dimulai dengan menentukan pengelompokan data dengan menentukan

himpunan dan atribut, berdasarkan data pada KOPERIA.

**Tabel 1. Himpunan Atribut**

Atribut	Kategori
Total kasus	
Jumlah Penghasilan	> 5 Juta 3-5 Juta
Jumlah Saldo	> 3 Juta 1-3 Juta < 1 Juta
Jumlah Pinjaman	11-20 Juta 6-10 Juta 1-5 Juta
Jangka Waktu	1-4 Bulan 5-8 Bulan 9-12 Bulan
Keperluan	Modal Usaha Membeli Barang Lain-lain

**Table 2 data training Set Nasabah**

No	No Agt	Nama Anggota	Jumlah Penghasilan	Jumlah Saldo	Jumlah Pinjaman	Jangka Waktu (Bulan)	Keperluan	Status
1	A-004	Ny Syofina	7,500,000	2,996,400	10,000,000	10	Modal usaha	Layak
2	A-001	Djoko Subadjo	7,500,000	4,017,000	15,000,000	10	Modal usaha	Layak
3	A-016	Ny Sainah Marjono	7,500,000	2,898,000	6,000,000	10	Membeli barang	Layak
4	A-018	Des Sutigno	9,300,000	2,833,000	15,000,000	10	Modal usaha	Layak
5	A-024	Sugro Kartorano	8,000,000	2,874,400	6,000,000	10	Membeli barang	Layak
6	A-038	Ny Suparniah	7,600,000	2,856,500	8,000,000	5	Membeli barang	Layak
7	A-089	Bagian ButarButar	8,300,000	2,878,400	6,000,000	6	Membeli barang	Layak
8	A-091	Ny Sadisem	8,000,000	2,821,400	5,000,000	5	Lain-lain	Tidak Layak
9	A-094	Sinin Hidayat	7,400,000	2,870,400	10,000,000	10	Modal usaha	Layak
10	A-095	Moch Soepri	7,400,000	2,882,400	5,000,000	5	Lain-lain	Tidak Layak
11	A-277	Suryanto	3,500,000	2,668,000	15,000,000	8	Modal usaha	Layak
12	A-178	Soekarto	6,500,000	2,685,000	4,000,000	5	Membeli barang	Tidak Layak
13	A-179	Siswut Hartono	7,500,000	3,510,000	8,000,000	10	Modal usaha	Layak
14	A-180	Ny Yohana Kristiani	7,500,000	2,847,000	10,000,000	5	Modal usaha	Layak
15	A-190	Ny Hj Nurana	8,000,000	3,422,500	13,000,000	10	Modal usaha	Layak
16	A-193	Ny Tati Sukarni	8,300,000	2,884,000	4,000,000	5	Lain-lain	Tidak Layak
17	A-201	Sumaryono	6,500,000	2,894,000	5,000,000	10	Lain-lain	Tidak Layak
18	A-204	Ny Sarinah	6,600,000	3,297,000	10,000,000	5	Modal usaha	Layak
19	439	Surmanan	3,200,000	1,280,000	4,000,000	2	Lain-lain	Layak
20	A-327	Fachri Rachman	3,400,000	2,857,000	6,000,000	6	Lain-lain	Tidak Layak
21	A-213	Santoso	5,500,000	2,848,000	4,000,000	5	Lain-lain	Tidak Layak
22	A-215	Ny Kaniyah Suprpto	6,500,000	2,832,000	7,000,000	7	Lain-lain	Tidak Layak
23	A-219	Ny Maria P Suprpto	6,500,000	3,047,000	6,000,000	6	Lain-lain	Tidak Layak
24	A-221	Ny Karsan	4,150,000	880,000	5,000,000	5	Membeli barang	Tidak Layak
25	A-225	Ny Turyah	5,500,000	2,896,000	4,000,000	5	Lain-lain	Tidak Layak
26	A-227	Ny Soetini	6,500,000	2,876,500	8,000,000	8	Membeli barang	Layak
27	A-236	Ny Aisah Kusnati	6,500,000	2,768,000	10,000,000	8	Modal usaha	Layak
28	A-256	Eandang Suprpto	3,300,000	2,704,000	4,000,000	5	Lain-lain	Layak
29	A-240	Ny Tuti Padji S	6,600,000	966,000	5,000,000	5	Membeli barang	Tidak Layak
30	A-242	Rahmat	6,500,000	2,754,000	7,000,000	7	Lain-lain	Tidak Layak

Guna tercapainya akurasi untuk menentukan kelayakan kredit maka metode yang dipakai adalah Algoritma C4.5, sesuai dengan penelitian yang sudah dilakukan sebelumnya.

Evaluasi dan validasi hasil dalam penelitian ini adalah klasifikasi pada data *training* berdasarkan waktu dan proses yang dibutuhkan, tingkat akurasi serta jumlah data yang diklasifikasikan kedalam kelas YA/TIDAK. Untuk mengetahui evaluasi

dari kinerja model klasifikasi didasarkan pada banyaknya (*count*) ditabulasikan dalam sebuah hasil yang dikenal sebagai *confusion matrix*. (CRISP-DM) (Larose, 2005) terdiri dari tahap antara lain:

1. *business understading*, data *understanding*, data yang diambil berdasarkan laporan kredit tahun 2015 sebanyak 30 nasabah, berdasarkan jumlah penghasilan, jumlah saldo, jumlah pinjaman, jangka waktu dan keperluan. Berikut nama atribut, kategori, dan nilai (*rangeny*). Berikut rulanya nilainya:

2. *Tahapan data prepetion*

Pada table 2 menunjukkan data training set data nasabah:

**HASIL DAN PEMBAHASAN**

Tujuan dari penelitian ini menguji keakuratan analisa kredit, untuk kelayakan pengajuan kredit dengan menggunakan algoritma C4.5. data yang dianalisa untuk proses merupakan data nasabah KOPERIA.

**Pengujian Model**

1. *Entropy*

Menghitung nilai total kasus keseluruhan yang ada berdasarkan jumlah kasus yang layak dan tidak layak. Kemudian akan diproses dengan mencari nilai *Entropy* dari setiap atribut dapat ditunjukkan pada rumus (1).

$$Entropy(S) = \sum_{i=1}^n -p_i * \log_2 p_i \dots \dots \dots (1)$$

$$Entropy(S)_{total} = \left( \left( \frac{\text{total layak}}{\text{total kasus}} \right) * \log_2 \left( \frac{\text{total layak}}{\text{total kasus}} \right) \right) + \left( \left( \frac{\text{total tidak layak}}{\text{total kasus}} \right) * \log_2 \left( \frac{\text{total tidak layak}}{\text{total kasus}} \right) \right) \dots \dots \dots (2)$$

$$Entropy (S) total = ((-17/30) * \log_2 (17/30)) + ((-13/30) * \log_2 (13/30)) = 0,9871$$

Perhitungan untuk nilai *entropy* pada setiap himpunan dan atribut, dibawah ini merupakan proses perhitungannya dan Hasil perhitungan dapat dilihat pada table 3.

Menghitung nilai *entropy* untuk  $S_{>5Juta}, S_{3-5Juta}$  untuk Jumlah Penghasilan antara lain:

$$Entropy (> 5Juta) = ((-14/25) * \log_2 (14/25)) + ((-11/25) * \log_2 (11/25)) = 0,468 441 + 0,521 147 = 0,9896$$

$$Entropy (3-5Juta) = ((-3/5) * \log_2 (3/5)) + ((-2/5) * \log_2 (2/5)) = 0,442 179 + 0,528 771 = 0,971$$

2. *Information Gain*

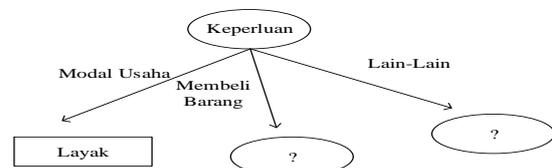
Perhitungan nilai *Gain* pada setiap himpunan atribut, dibawah ini proses perhitungannya. Menghitung nilai *Gain* dapat ditunjukkan pada rumus (3).

$$Gain (S, A) = Entropy(S) - \sum_{i=0}^n \frac{|S_i|}{|S|} * Entropy(S_i) \dots \dots \dots (3)$$

$$Gain (jumlah penghasilan) = 0,9871 - ((25/30*0,9896) + (5/30*0,971)) = 0,001$$

3. Proses pembuatan pohon

Dari hasil perhitungan pada table 3 dapat di ketahui nilai *Gain* yang terbesar yaitu 0,473 adalah himpunan Keperluan, maka hasil pohon keputusan sementara dapat dilihat pada gambar 1.



**Gambar 1 Pohon keputusan pada node 1**

Dengan memperhatikan pohon keputusan pada gambar 2 diketahui bahwasemua kasus sudah masuk dalam kelas, dengan demikian pohon keputusan pada gambar merupakan pohon keputusan terakhir yang terbentuk. Dihasilkan sejumlah aturan dalam pohon tersebut. Berikut ini aturan

yang dihasilkan pada pohon keputusan akhir sebagai berikut:

“Jika Keperluan Membeli Barang dan Jumlah Pinjaman 11-20 Juta Maka = Tidak Layak”

“Jika Keperluan Membeli Barang dan Jumlah Pinjaman 6-10 Juta Maka = Layak”

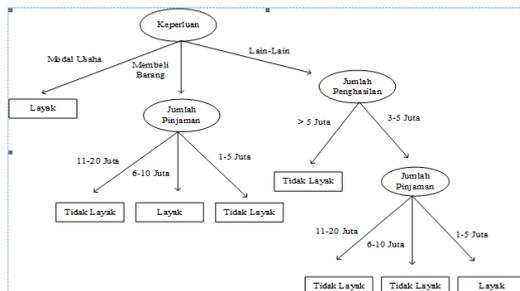
“Jika Keperluan Membeli Barang dan Jumlah Pinjaman 1-5 Juta Maka = Tidak Layak”

“Jika Keperluan Lain-lain dan Jumlah Penghasilan > 5 Juta Maka = Tidak Layak”

“Jika Keperluan Lain-lain dan Jumlah Penghasilan 3-5 Juta dan Jumlah Pinjaman 11-20 Juta Maka = Tidak Layak”

“Jika Keperluan Lain-lain dan Jumlah Penghasilan 3-5 Juta dan Jumlah Pinjaman 6-10 Juta Maka = Tidak Layak”.

“Jika Keperluan Lain-lain dan Jumlah Penghasilan 3-5 Juta dan Jumlah Pinjaman 1-5 Juta Maka = “Layak”.



**Gambar 2. Pohon Keputusan Akhir**

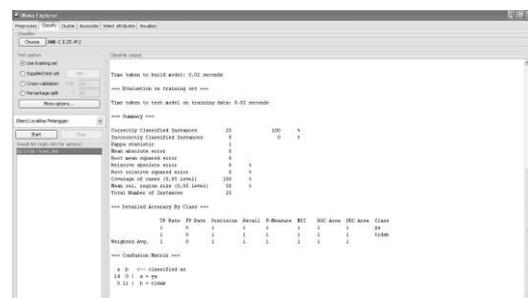
Hasil klasifikasi pada data sampling atribut keperluan sebagai *root* pada *decision tree*, sedangkan atribut lainnya sebagai *child node*. Dari data sampling 30 *record* dihasilkan jumlah aturan yang terbentuk sebanyak 7 aturan.

Tahap selanjutnya proses pengolahan data dari hasil yang diperoleh dari konsep algoritma C4.5. sebelum di proses dengan aplikasi WEKA data di pecah menjadi 2 bagian, yaitu pertama sebagai data *training* sebanyak 75% data testing sebanyak 35% hal ini supaya dapat terbentuk model dengan menggunakan data *training*, selanjutnya data yang terbentuk akan diujikan kembali menggunakan data *testing*.

**Table 3 Nilai Perhitungan Entropy dan Gain**

Atribut	Kategori	Banyak Kasus	Layak	Tidak Layak	Entropy	Gain
Total kasus		30	17	13	0,9871	
Jml Penghasilan						0,001
	> 5 Juta	25	14	11	0,9896	
	3-5 Juta	5	3	2	0,971	
Jml Saldo						0,11
	> 3 juta	5	4	1	0,7219	
	1-3 Juta	23	13	10	0,9877	
	< 1 Juta	2	0	2	0	
Jml Pinjaman						0,318
	11-20 Juta	4	4	0	0	
	6-10 Juta	15	11	4	0,8366	
	1-5 Juta	11	2	9	0,684	
Jangka Waktu						0,189
	1-4 Bulan	1	1	0	0	
	5-8 Bulan	20	8	12	0,971	
	9-12 Bulan	9	8	1	0,5033	
Keperluan						0,473
	Modal Usaha	10	10	0	0	
	Membeli Barang	8	5	3	0,9544	
	Lain-lain	12	2	10	0,650	

Hasil klasifikasi data *training* sebanyak 65% dengan menggunakan WEKA dapat dilihat pada gambar 3.



**Gambar 3. Hasil Klasifikasi data Training Set menggunakan WEKA.**

Pada saat menggunakan data *training* dengan menggunakan WEKA 3.7.7 waktu yang dibutuhkan untuk menguji

model pada *training* ini adalah 0.02 second dalam mengklasifikasikan diaplikasi weka 3.7.7. dari *confusion matrix* dapat dilihat ada 13 rekords data diklasifikasikan sebagai *class* YA, 10 rekords data yang diklasifikasikan sebagai *class* TIDAK dan 2 record diklasifikasikan kedalam *class* yang tidak sesuai yaitu data diduga sebagai *class* YA ternyata adalah *class* TIDAK. Hasil akurasi detail dapat dilihat pada gambar 4.

--- Detailed Accuracy By Class ---

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	1	0	1	1	1	1	1	1	YA
	1	0	1	1	1	1	1	1	TIDAK
Weighted Avg.	1	0	1	1	1	1	1	1	

**Gambar 4. Hasil Akurasi Detail oleh Class Training**

Tingkat positif benar (TP) *Rate* adalah proporsi contoh yang diklasifikasikan sebagai kelas x, diantara semua contoh yang benar-benar memiliki kelas x yang bernilai sama dengan *recall*.

Untuk mencari nilai TP *rate* pada *class* YA, pada rumus 4:

$$\text{Positive predictive value} = \frac{14}{14 + 0} = 1 \quad (4)$$

Untuk mencari nilai FN sama pada *class* Tidak ,pada rumus 5:

$$\text{False Negatif} = \frac{11}{11 + 0} = 1 \quad (5)$$

Tingkat positif salah (FP) *rate* adalah proporsi yang diklasifikasikan sebagai kelas x, tetapi masuk kedalam kelas yang berbeda diantara semua contoh yang bukan dari kelas x.

Untuk mencari nilai FP sama pada *class* Ya, berikut rumusnya (6) :

$$\text{False Positive} = \frac{0}{0 + 14} = 0 \quad (6)$$

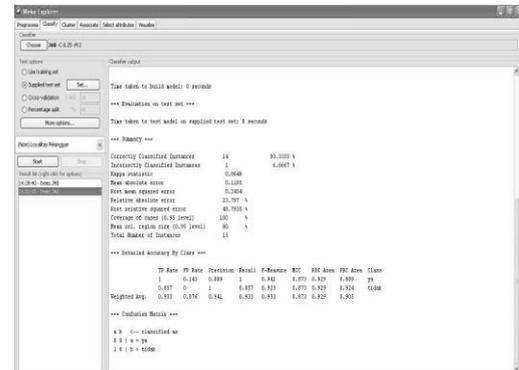
Untuk mencari nilai FN sama pada *class* Tidak, berikut rumusnya (7):

$$\text{False Negative} = \frac{0}{0 + 11} = 0 \quad (7)$$

Untuk mencari *precision class* YA, berikut rumusnya (8):

$$\frac{14}{14 + 0} = 1 \quad (8)$$

Proses selanjutnya yaitu klasifikasi menggunakan data *testing* sebanyak 35% dengan menggunakan weka. Hasil klasifikasi menggunakan data *testing* pada gambar 5.



**Gambar 5. Hasil Klasifikasi data testing**

Waktu yang dibutuhkan untuk menguji model *training* ini 0 *second* dalam mengklasifikasikan dari *confusion matrix* dapat dilihat ada 8 *rekords* data diklasifikasikan sebagai *class* YA, 6 *rekords* data yang diklasifikasikan sebagai *class* TIDAK dan 1 *record* diklasifikasikan kedalam *class* yang tidak sesuai yaitu data diduga sebagai *class* YA ternyata adalah *class* TIDAK.

Tingkat positif benar (TP) *Rate* adalah proporsi contoh yang diklasifikasikan sebagai kelas x, diantara semua contoh yang benar-benar memiliki kelas x yang bernilai sama dengan *recall*.

Untuk mencari nilai TP *rate* pada *class* YA berikut rumusnya (9).

$$\text{Positive predictive value} = \frac{8}{8 + 0} = 1 \quad (9)$$

Untuk mencari nilai FN sama pada *class* Tidak berikut rumusnya (10):

$$\text{False Negative} = \frac{6}{6 + 8} = 0.143 \quad (10)$$

Tingkat positif salah (FP) *rate* adalah proporsi yang diklasifikasikan sebagai kelas x, tetapi masuk kedalam kelas yang berbeda diantara semua contoh yang bukan dari kelas x.

Untuk mencari nilai FP sama pada *class* Ya berikut rumusnya (11):

$$\text{false Positive} = \frac{6}{8 + 1} = 0.143 \quad (11)$$

Untuk mencari nilai FN sama pada *class* Tidak berikut rumusnya (12):

$$\text{false Negative} = \frac{0}{0 + 6} = 0 \quad (12)$$

Untuk mencari *precision class* YA berikut rumusnya (13).

$$\frac{9}{8 + 1} = 0.889 \quad (13)$$

Untuk mencari *precision class* Tidak, berikut rumusnya (14).

$$\frac{6}{6 + 0} = 1 \quad (14)$$

*F-measure* yang merupakan ukuran gabungan dari *precision* dan *recal*. proses perhitungannya pada rumus (15).

$$F - \text{Measure} = \frac{2 * \text{Precision} * \text{recal}}{\text{Precision} + \text{Recall}} \quad (15)$$

Untuk mencari *F-Measure class* Ya pada rumus (16):

$$\frac{2 * 1 * 0.867}{1 * 0.867} = 0.929 \quad (16)$$

Untuk mencari *F-measure class* TIDAK pada rumus (17).

$$\frac{2 * 0.833 * 1}{0.833 + 1} = 0.909 \quad (17)$$

$$\frac{2 * 0.833 * 1}{0.833 + 1} = 0.909$$

Tahap selanjutnya mengukur evaluasi untuk mengetahui evaluasi dari kinerja model klasifikasi. Pada weka *classifier* hasil klasifikasi yang diperoleh akan disertakan dengan beberapa alat ukur sebagai bukti terjadinya hasil proses pada klasifikasi tersebut yang telah tersedia didalamnya, salah satu alat ukur tersebut adalah *confusion matrix*. Berikut ini adalah hasil kesimpulan *confusion matrik* yang terdaat pada aplikasi WEKA dari hasil yang diperoleh dari data *training*, dapat dilihat pada gambar 6.

=== Confusion Matrix ===

```
a b <-- classified as
14 0 | a = ya
0 11 | b = tidak
```

Gambar 6. Hasil Confusion Matrix

Dari informasi tersebut, kemudian akan dilakukan proses perhitungan rata-rata persentasiakurasi keberhasilan dan *error rate* pada *confusion matrix* data training dapat ditunjukkan pada perhitungan (18).

$$\text{Akurasi} = \frac{\text{banyaknya prediksi yang benar}}{\text{total banyaknya prediksi}} = \frac{f_{11} + f_{00}}{f_{11} + f_{10} + f_{01} + f_{00}} \quad (18)$$

$$\text{Error Rate} = \frac{\text{banyaknya prediksi yang salah}}{\text{total banyaknya prediksi yang salah}} = \frac{f_{10} + f_{01}}{f_{11} + f_{10} + f_{01} + f_{00}}$$

$$\text{Akurasi} = \frac{14 + 11}{14 + 0 + 0 + 11} = 1$$

Persentase Akurasi = 1x 100% = 100%

$$\text{Error Rate} = \frac{0 + 0}{14 + 0 + 0 + 11} = 0$$

Persentase Error= 0 x 100%= 0%

Proses selanjutnya adalah hasil kesimpulan *confusion matrix* yang terdapat pada aplikasi WEKA dari hasil yang diperoleh menggunakan data nasabah yang diambil merupakan data *testing* dapat dilihat pada gambar 7.

=== Confusion Matrix ===

```
a b <-- classified as
9 0 | a = ya
2 4 | b = tidak
```

Gambar 7. Hasil *confusion matrix* data *testing*

$$\text{Akurasi} = \frac{\text{banyaknya prediksi yang benar}}{\text{total banyaknya prediksi}} = \frac{f_{11} + f_{00}}{f_{11} + f_{10} + f_{01} + f_{00}}$$

$$\text{Error Rate} = \frac{\text{Banyaknya prediksi yang salah}}{\text{Total banyaknya prediksi}} = \frac{f_{10} + f_{01}}{f_{11} + f_{10} + f_{01} + f_{00}}$$

$$\text{Akurasi} = \frac{9 + 4}{9 + 0 + 2 + 4} = 0.86667$$



Tampilan hasil data keseluruhan yang sudah di proses dengan Algoritma C4.5, yang nantinya sudah menjadi keputusan dalam penentuan kelayakan kredit.

### KESIMPULAN

Berdasarkan analisa penggunaan data mining dengan algoritma C4.5 terbukti akurat dalam penentuan pengajuan kelayakan kredit. Hal ini dibuktikan dengan hasil evaluasi penenelitian bahwa algoritma C4.5 mendapatkan nilai akurasi *data training* serta *data testing* menggunakan algoritma C4.5, kedua proses tersebut rata-rata 93%, dan *error rate* 6.6%

### DAFTAR PUSTAKA

- [1] Andi. (2016). *Pemrograman PHP dan MySQL*. . Yogyakarta: MADCOMS.
- [2] Firmansyah. (2011). *Penerapan Algoritma Klasifikasi C4.5 Untuk Penentuan Kelayakan Pemberian Kredit Koperasi*. Jakarta: STMIK Nusa Mandiri.
- [3] Haryati, Siska, Sudarsono, A., & Suryana, E. (2015). Implementasi Data Mining Untuk Memprediksi Masa Studi Mahasiswa Menggunakan Algoritma C4.5 (Studi Kasus: Universitas Dehasen Bengkulu). *Jurnal Media Infotama*.
- [4] Kusriani, E. T. (2009). *Algoritma Data Mining*. Yogyakarta: CV Andi Offset.
- [5] Larose, D.T. (2005). *Discovering Knowledge in Data*.
- [6] Masripah, S. (2015). Evaluasi Penentuan Kelayakan Pemberian Kredit Koperasi Syariah Menggunakan Algoritma Klasifikasi C4.5. *Jurnal Pilar Nusa Mandiri*, XI.
- [7] Munawar. (2005). *Pemodelan Visual Dengan UML*. Yogyakarta: Graha Ilmu.
- [8] Rusito, M. T. (2016). Implementasi Metode Decision Tree Dan Algoritma C4.5 Untuk Klasifikasi Data Nasabah Bank. *Infokam*, Vol.1.
- [9] Santoso, T. B. (2014). Analisa dan Penerapan Metode C4.5 untuk Prediksi Loyalitas Pelanggan. *Jurnal Ilmiah Fakultas Teknik - LIMIT'S Vol.10 (1) Hal.22-31*.