

# KAJIAN PENERAPAN METODE KLASIFIKASI DATA MINING ALGORITMA C4.5 UNTUK PREDIKSI KELAYAKAN KREDIT PADA BANK MAYAPADA JAKARTA

Nandang Iriadi<sup>1</sup>, Nia Nuraeni<sup>2</sup>

**Abstract** — *The banking industry has developed quite rapidly, both in terms of volume of business, mobilize public funds or credit. Data mining of the loan has great potential to explore the hidden patterns within a dataset of loans including loans domain. C4.5 classification algorithm is the most simple, easy to implement. However, the algorithm C4.5 still has weaknesses in handling high-dimensional data in. This research aims to implement the algorithm C4.5 with the selection of attributes so as to reduce the dimensionality of the data, and identify features in the data set with C4.5 algorithm method. From this research, conducted models created with C4.5 algorithm itself already has good accuracy that is equal to 83.67% with the selection process by the algorithm C4.5 attributes.*

**Intisari** — *Industri perbankan mengalami perkembangan yang cukup pesat, baik dari sisi volume usaha, mobilisasi dana masyarakat maupun pemberian kredit. Data mining mengenai pinjaman memiliki potensial besar untuk menjelajahi bagian pola yang tersembunyi dalam suatu dataset dari domain pinjaman termasuk pinjaman kredit. Algoritma C4.5 merupakan pengklasifikasian yang paling sederhana, mudah diimplementasikan. Namun, Algoritma C4.5 masih memiliki kelemahan dalam menangani data dalam dimensi tinggi. Penelitian ini bertujuan untuk menerapkan algoritma C4.5 dengan seleksi atribut sehingga dapat mengurangi dimensi dari data, serta mengidentifikasi fitur dalam kumpulan data dengan metode algoritma C4.5. Dari penelitian ini yang dilakukan model yang terbentuk dengan algoritma C4.5 sendiri sudah memiliki akurasi yang baik yaitu sebesar 83.67% dengan proses seleksi atribut oleh algoritma C4.5.*

**Kata kunci** — *Data mining, Algoritma C4.5, kelayakan kredit, Decision Tree*

<sup>1</sup> Program Studi Teknik Komputer AMIK BSI Jakarta, Jl. Fatmawati No. 24 Pondok Labu DKI Jakarta, Telp: 021-7500282, 7500680 Fax: 021-7513790 Email : [ndandang.ndi@bsi.ac.id](mailto:ndandang.ndi@bsi.ac.id)

<sup>2</sup> Program Studi Manajemen Informatika AMIK BSI Jakarta, Jl. Kramat Raya No. 18 DKI Jakarta, Telp: 021- 021-3100413, Fax: 021-3144869 Email : [nia.nne@bsi.ac.id](mailto:nia.nne@bsi.ac.id)

## I. PENDAHULUAN

Bank Mayapada Mitra Usaha Cabang PGC merupakan Bank konvensional yang memberikan pinjaman kredit kepada calon debitur yang memiliki usaha mikro, kecil dan menengah khususnya pengusaha mikro yang berada dilingkungan PGC (Pusat Grosir Cililitan). Usaha Mikro sebagaimana dimaksud menurut keputusan menteri keuangan No.40/KMK.06/2003 tanggal 29 Januari 2003, yaitu usaha produktif milik keluarga atau perorangan Warga Negara Indonesia dan memiliki hasil penjualan paling banyak Rp. 100.000.000,00 (Seratus Juta Rupiah) per tahun. Usaha mikro dapat mengajukan kredit kepada Bank paling banyak Rp. 50.000.000 (Lima Puluh Juta Rupiah). Dilihat dari kepentingan perbankan, usaha mikro adalah suatu segmen pasar yang cukup potensial untuk dilayani dalam upaya meningkatkan fungsi intermediasinya karena usaha mikro mempunyai karakteristik positif dan unik yang tidak selalu dimiliki oleh usaha non mikro. Namun demikian disadari sepenuhnya bahwa masih banyak usaha mikro yang sulit memperoleh layanan kredit perbankan karena berbagai kendala baik pada sisi usaha mikro maupun pada sisi perbankan sendiri.

Dalam proses pemberian kredit selama ini, khususnya pemberian kredit usaha mikro yang dilakukan oleh Bank Mayapada meskipun melalui analisa kredit masih saja ada permasalahan yang timbul diantaranya, para calon debitur melakukan segala macam cara agar kreditnya disetujui oleh pihak bank. Hal ini yang menyebabkan tingkat kredit macet juga meningkat. Penyebabnya antara lain kurang akuratnya *credit officer* dalam memberikan analisisnya, pihak *account officer* yang kurang tepat dalam mencari calon debitur karena dikejar oleh target perusahaan. Tujuan penelitian ini adalah :

- 1) Menerapkan algoritma C4.5 untuk meningkatkan keakuratan analisa kredit pada Bank Mayapada Mitra Usaha Cabang PGC.
- 2) Penggunaan algoritma algoritma C4.5 serta mengoptimalkan dan menganalisa atribut faktor usaha sebagai parameter untuk penentuan kelayakan kredit, diantaranya: jenis usaha, status tempat usaha, lama usaha, sistem penjualan, sistem pembelian, omzet per bulan, *gross profit margin*, *repayment capacity*, fasilitas dan kolek nasabah.

II. KAJIAN LITERATUR

A. Data Mining

Data Mining adalah proses menemukan korelasi baru yang bermakna, pola dan tren dengan memilah-milah sejumlah besar data yang tersimpan dalam repositori, menggunakan teknologi penalaran pola serta teknik-teknik statistik dan matematika[4]. Pada prosesnya *data mining* akan mengekstrak informasi yang berharga dengan cara menganalisis adanya pola-pola ataupun hubungan keterkaitan tertentu dari data-data yang berukuran besar. Dengan adanya data mining dapat menemukan pengetahuan yang dapat dipergunakan oleh para praktisi kesehatan agar dapat meningkatkan kualitas pelayanan. Pengetahuan yang ditemukan dapat juga digunakan oleh praktisi medis untuk mengurangi jumlah efek samping dari obat, untuk menyarankan lebih murah dalam terapi yang setara dengan berbagai alternatif. Proses *mining* data lebih dari sekedar analisa data yang meliputi pengklasifikasian, klastering, asosiasi dan prediksi [1].

B. Algoritma C4.5

C4.5 adalah bagian dari algoritma untuk klasifikasi dalam pembelajaran *machine learning* dan data mining. C4.5 merupakan algoritma yang cocok digunakan untuk masalah klasifikasi pada *machine learning* dan *data mining* [6]. Dalam pembuatan pohon keputusan, setiap algoritma menerapkan ukuran pemilihan atribut yang berbeda-beda. Ukuran pemilihan atribut merupakan ukuran yang digunakan dalam menentukan kriteria yang terbaik untuk mengelompokkan *tuple*. Ukuran pemilihan atribut ini juga disebut sebagai *splitting rules* karena menentukan bagaimana data akan dipisahkan kesetiap cabang. *Information gain* diciptakan oleh Claude Shannon dengan mempelajari nilai informasi dari data<sup>a</sup>, dan menggunakan nilai tersebut sebagai acuan dalam menentukan atribut yang akan digunakan dalam menyusun pohon keputusan[3]. Konsep dasar dari *decision tree* adalah mengubah data menjadi pohon keputusan dengan aturannya. Atribut yang dipilih akan menghasilkan partisi dengan data yang lebih seragam, dan dapat menghasilkan pohon keputusan yang sesederhana mungkin dengan perulangan yang sedikit. Sebuah pohon keputusan terdiri dari sekumpulan aturan yang bertujuan untuk membagi sejumlah populasi yang heterogen menjadi lebih kecil dan lebih homogen dengan memperhatikan variabel tujuannya [5]. Algoritma C4.5 menggunakan konsep *information gain* atau *entropy reduction* untuk memilih pembagian yang optimal [4]. Membuat penelitian untuk membandingkan beberapa algoritma seperti Regresi Linier, Neural Network, Support Vector Machine, Case Base Reasoning, Rule Based Fuzzy Neural Network dan Decision Tree. Semua model algoritma tadi digunakan untuk menganalisa persetujuan pinjaman dalam bentuk kredit. Dari hasil penelitian didapatkan bahwa *Decision Tree* terbukti mempunyai akurasi tertinggi dalam menentukan keputusan dibandingkan algoritma lain. [8].

Tahapan dalam membuat sebuah pohon keputusan dengan algoritma C4.5 [2]. yaitu :

- 1) Mempersiapkan data *training*, dapat diambil dari data histori yang pernah terjadi sebelumnya dan sudah dikelompokkan dalam kelas-kelas tertentu.
- 2) Menentukan akar dari pohon dengan menghitung nilai *gain* yang tertinggi dari masing-masing atribut atau berdasarkan nilai *index entropy* terendah. Sebelumnya dihitung terlebih dahulu nilai *index entropy*, dengan rumus :

$$Entropy(i) = - \sum_{j=1}^m f(i,j) \cdot \log_2 f(i,j)$$

Keterangan :

- i = himpunan kasus
- m = jumlah partisi i
- f(i,j) = proporsi j terhadap i

- 3) Hitung nilai *gain* dengan rumus :

$$Entropy\ split = - \sum_{i=1}^p \frac{n_i}{n} \cdot IE(i)$$

Keterangan :

- p = jumlah partisi atribut
  - ni = proporsi ni terhadap i
  - n = jumlah kasus dalam n
- 4) Ulangi langkah ke-2 hingga semua *record* terpartisi. Proses partisi pohon keputusan akan berhenti disaat :  
Semua tupel dalam *record* dalam simpul m mendapat kelas yang sama. Tidak ada atribut dalam *record* yang dipartisi lagi
  - 5) Tidak ada *record* di dalam cabang yang kosong. Algoritma C4.5 menggunakan konsep *information gain* atau *entropy reduction* untuk memilih pembagian yang optimal [4]. Tahapan dalam membuat sebuah pohon keputusan dengan algoritma C4.5 [2] yaitu :
    - a. Mempersiapkan data *training*, dapat diambil dari data histori yang pernah terjadi sebelumnya dan sudah dikelompokkan dalam kelas-kelas tertentu.
    - b. Atribut atau berdasarkan nilai *index entropy* terendah. Sebelumnya dihitung terlebih dahulu nilai *index entropy*, dengan rumus :

$$Entropy(i) = - \sum_{j=1}^m f(i,j) \cdot \log_2 f(i,j)$$

Keterangan :

- i = himpunan kasus
  - m = jumlah partisi i
  - f(i,j) = proporsi j terhadap i
- c. Hitung nilai *gain* dengan rumus :



$$Entropy(i) = - \sum_{j=1}^m f(i,j) \cdot \log_2 f(i,j)$$

Keterangan :

p = jumlah partisi atribut

ni = proporsi ni terhadap i

n = jumlah kasus dalam n

8. Ulangi langkah ke-2 hingga semua *record* terpartisi.

Proses partisi pohon keputusan akan berhenti disaat :

- Semua tupel dalam *record* dalam simpul m mendapat kelas yang sama.
- Tidak ada atribut dalam record yang dipartisi lagi.
- Tidak ada record di dalam cabang yang kosong.

### III. METODE PENELITIAN

Penelitian yang digunakan adalah penelitian *Experiment*. Penelitian eksperimen melibatkan penyelidikan perlakuan parameter atau variabel tergantung pada penelitiannya dan menggunakan tes yang dikendalikan oleh si peneliti itu sendiri.

#### A. Penelitian *Experiment*

Pada metode penelitian eksperimen, digunakan model proses CRISP-DM (*Cross- Standard Industry Process for Data Mining*) yang terdiri dari 6 tahap [4]:

- Business Understanding*
- Data Understanding*
- Data Preparation*
- Modelling*
- Evaluation*
- Deployment*

#### B. *Business Understanding*

Berdasarkan data nasabah kredit tahun 2014, terlihat bahwa nasabah dengan status kreditnya macet lebih banyak dari nasabah yang status kreditnya lancar, hal ini menjadi sebuah masalah dan kendala bagi Bank Mayapada Mitra Usaha cab PGC khususnya karna akan berakibat pada kurang akuratnya analisa kredit. Dalam beberapa penelitian telah dilakukan proses analisa kredit dengan menggunakan Algoritma C4.5, dengan tujuan analisa yang dilakukan akan lebih akurat.

#### C. *Data Understanding*

Data yang didapatkan dari Bank Mayapada Mitra Usaha Cab PGC dan sekitarnya (plasma) adalah data kredit nasabah tahun 2014. Atribut atau variabel yang ada sebanyak 22 atribut (data lengkapnya bisa dilihat dilampiran). Setelah dilakukan proses data *preparation*, atribut atau variabel yang digunakan terdiri dari 10 atribut yang terdapat dalam data status kredit nasabah. Variabel prediktor yaitu jenis usaha, status tempat usaha, lama usaha, sistem penjualan, sistem pembelian, omset per bulan, *gross profit margin*, *repayment capacity*, dan fasilitas. Sedangkan variabel tujuan adalah kolek nasabah (yang menunjukkan status kredit lancar atau macet).

#### D. *Data Preparation*.

Pada tahapan ini data sebanyak 300 dan atribut yang terdiri dari 22 atribut, akan dilakukan beberapa penyeleksian untuk menghasilkan data yang dibutuhkan, tahapannya yaitu:

- Data Cleaning* untuk membersihkan nilai yang kosong atau tuple yang kosong.
- Data Integration* yang berfungsi menyatukan tempat penyimpanan yang berbeda kedalam satu data.
- Data Reduction*, jumlah atribut yang digunakan terlalu banyak dan tidak semua atribut menjadi syarat atas atribut penentu sehingga yang digunakan hanya 10 Atribut

#### E. *Modelling*

Pada tahap ini, data *preparation* yang telah didapatkan akan dilakukan pemodelan Algoritma C4.5, dimana dalam probabilitas itu dilakukan perhitungan manual, kemudian dimasukan kedalam *tools* rapidminer dengan menggunakan algoritma C4.5

### IV. HASIL DAN PEMBAHASAN

#### 1. Pengolahan Awal Data

Jumlah data yang diperoleh dari data sekunder yaitu sebanyak 300 yang terdiri dari 112 data yang dinyatakan pembayaran kreditnya Lancar dan 188 yang pembayarannya macet. Dan pada data ini memiliki atribut atau variabel sebanyak 10 atribut yang terdiri dari 9 atribut *predictor* dan 1 atribut tujuan. Atribut *predictor* yang digunakan dalam data kredit adalah jenis usaha, Status tempat usaha, lama usaha, *system penjualan*, sistem pembelian, omset perbulan, *cross profit margin*, *Repayment Capacity*, dan *Fasilitas*. Sedangkan atribut terakhir adalah sebagai *output* berupa keputusan yang digunakan untuk membedakan hasil prediksi, yaitu "Lancar" untuk Nasabah yang lancar dalam pembayaran pinjamannya dan "macet" untuk nasabah yang tidak lancar dalam pinjamannya.

Tabel 1. Attribute Data Kredit Nasabah Bank Mayapada Mitra Usaha Cab PGC

Jenis Usaha	Status Tempat Usaha	Lama Usaha	Sistem Penjualan
Penjualan Mobil Second	Kontrak	3 - 5 Tahun	Cash $\geq$ 80 % dan < 100 %, Kredit $\geq$ 10 % dan $\leq$ 20 %
Toko Obat dan Kontrakan	Kontrak	> 5 Tahun	Cash $\geq$ 80 % dan < 100 %, Kredit $\geq$ 10 % dan $\leq$ 20 %
Penggilingan batu	Milik Sendiri	> 5 Tahun	Cash $\geq$ 80 % dan < 100 %, Kredit $\geq$ 10 % dan $\leq$ 20 %

Sumber: Hasil Penelitian (2015)

#### 2. Model Algoritma C4.5

Penelitian ini bertujuan untuk mengembangkan model yang sudah terbentuk dengan algoritma C4.5. Data dianalisa dengan menggunakan algoritma C4.5.

1) **Pengujian Model**

Pada tahap ini dilakukan pengujian model yaitu menghitung dan mendapatkan rule-rule yang ada pada model algoritma C4.5 yang diusulkan. Dimana dalam penelitian ini bertujuan untuk mengembangkan model yang sudah terbentuk dengan algoritma C4.5. Data dianalisa dengan menggunakan algoritma C4.5.

Metode pengolahan algoritma C4.5, tahapan yang akan dilakukan adalah sebagai berikut:

Menghitung jumlah nasabah kredit yang memprediksikan pembayaran kreditnya lancar dan yang pembayaran kreditnya macet serta Entropy dari semua kasus yang dibagi berdasarkan atribut di tabel 1 dengan jumlah Nasabah 300, dengan jumlah kelas Lancar (112 record) dan macet (118 record). Baris total entropy dihitung berdasarkan data training sebagai berikut:

$$\text{Entropy}(\text{total}) = \left( -\frac{112}{300} * \log_2 \left( \frac{112}{300} \right) \right) + \left( -\frac{118}{300} * \log_2 \left( \frac{118}{300} \right) \right) = 0.327$$

Kemudian hitung nilai Entropy pada masing-masing atribut, sebagai contoh dibawah ini menghitung nilai Entropy untuk atribut Status tempat usaha dengan jumlah Kontrak 144 dengan kelas Lancar (34) dan Macet (110):

$$= \left( -\frac{34}{144} * \log_2 \left( \frac{34}{144} \right) \right) + \left( -\frac{110}{144} * \log_2 \left( \frac{110}{144} \right) \right) = 0.058$$

Milik Sendiri 156 dengan kelas Lancar (78) dan Macet (78)

$$= \left( -\frac{78}{156} * \log_2 \left( \frac{78}{156} \right) \right) + \left( -\frac{78}{156} * \log_2 \left( \frac{78}{156} \right) \right) = 0.3$$

1) Kemudian hitung gain untuk atribut diatas sebagai berikut

$$\text{gain} = 0,327 - \left( \left( \frac{144}{300} \right) * 0,58 \right) + \left( \frac{156}{300} * 0,3 \right)$$

$$G = -0,107$$

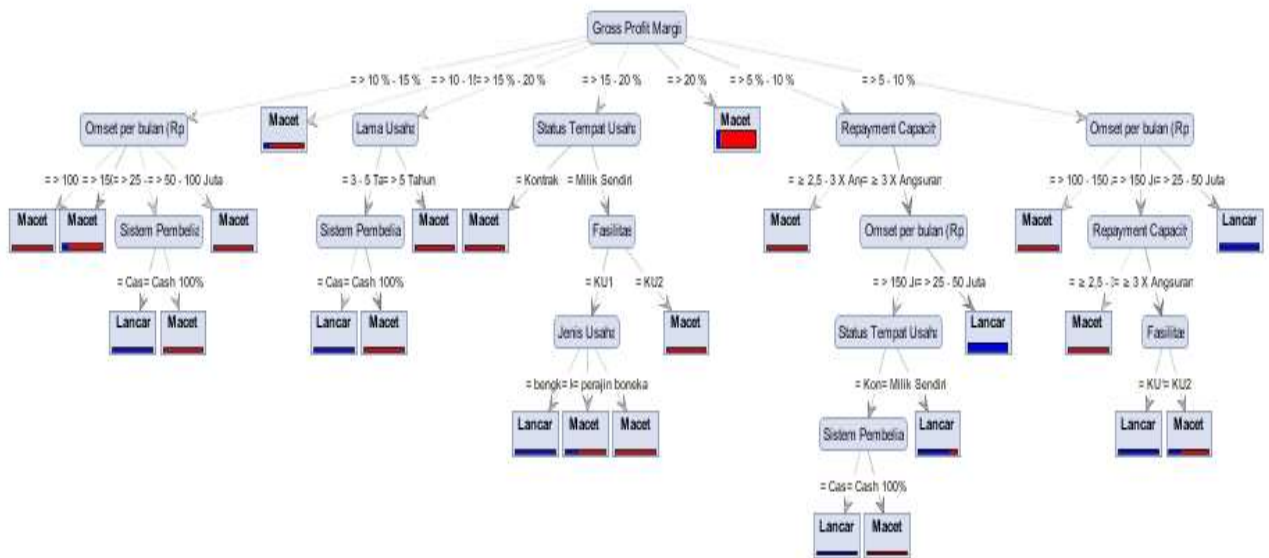
Milik Sendiri 156 dengan kelas Lancar (78) dan Macet (78)

$$= \left( -\frac{78}{156} * \log_2 \left( \frac{78}{156} \right) \right) + \left( -\frac{78}{156} * \log_2 \left( \frac{78}{156} \right) \right) = 0.3$$

1) Kemudian hitung gain untuk atribut diatas sebagai berikut

$$\text{gain} = 0,327 - \left( \left( \frac{144}{300} \right) * 0,58 \right) + \left( \frac{156}{300} * 0,3 \right)$$

$$G = -0,107$$



Sumber: Hasil Penelitian (2015)

Gambar 1. Pohon Keputusan C4.5



Tujuan utama dari menganalisis data dengan menggunakan algoritma *Decision Tree* ini adalah ingin mendapatkan rule yang akan dimanfaatkan untuk pengambilan keputusan pada data baru [5]. Adapun rule yang didapat dari Gambar 1 adalah:

1. R1: IF Gros Fropit > 10%-15% AND Omzet perbulan >100-150 juta AND >15-25 juta THEN class = Macet
2. R2 : IF Gros Fropit > 10%-15% AND Omzet perbulan >50 juta AND Sistem Pembelian =Cash THEN class = Lancar
3. R3: IF Gros Fropit > 15%-20% AND Lama Usaha 3-5 Tahun AND Sistem Pembelian =Cash THEN class = Lancar
4. R4: IF Gros Fropit > 15%-20% AND Lama Usaha > 5 Tahun THEN class = Macet
5. R5: IF Gros Fropit > 5%-10% AND Repayment Capacity > 2.5-3 kali angsuran THEN class = Macet
6. R6: IF Gros Fropit > 5%-10% AND Repayment Capacity > 3 kali angsuran AND omzet perbulan > 150 juta AND Status Tempat Usaha =kontrak AND Sistem pembelian=Cash THEN class = Lancar
7. R7 : IF Gros Fropit > 5%-10% AND Repayment Capacity > 3 kali angsuran AND omzet perbulan > 150 juta AND Status Tempat Usaha =milik sendiri THEN class = Lancar
8. R8 : IF Gros Fropit > 5%-10% AND omzet perbulan > 100-150 juta THEN class =Macet
9. R9: IF Gros Fropit > 5%-10% AND omzet perbulan > 150 juta AND Repayment Capacity >2.5-3 kali angsuran THEN class =Macet
10. R10: IF Gros Fropit > 5%-10% AND omzet perbulan > 150 juta AND Repayment
11. R11: IF Gros Fropit > 5%-10% AND omzet perbulan > 150 juta AND Repayment Capacity >3 kali angsuran AND Fasilitas = KU2 THEN class =Macet

Metode klasifikasi bisa dievaluasi berdasarkan beberapa kriteria seperti tingkat akurasi, kecepatan, kehandalan, skalabilitas, dan interpretabilitas [7]. Hasil pengujian model yang dilakukan untuk mengukur tingkat akurasi dan AUC (*Area Under Curve*) dari hasil diagnosa dengan metode *cross validation*. Hasil dari uji coba yang dilakukan yaitu untuk menghasilkan nilai *accuracy* dan nilai AUC (*Area Under Curve*).

## 2) Evaluasi model dengan Confusion matrix

Model *Confusion matrix* akan membentuk matrix yang terdiri dari *true positif* atau tupel positif dan *true negatif* atau tupel negatif. Berikut dibawah ini merupakan hasil dari *Confusion matrix* pada algoritma C4.5:

Tabel 1. Gambar Hasil *Accuracy* C4.5

accuracy: 83.67% +/- 4.58% (mikro: 83.67%)			
	true Lancar	true Macet	class precision
pred. Lancar	88	25	77.88%
pred. Macet	24	163	87.17%
class recall	78.57%	86.70%	

Sumber: Hasil Penelitian (2015)

Jumlah *True Positive* (TP) adalah 88 Record, *False Negatif* (FN) sebanyak 25, *True Negatif* (TN) sebanyak 24 dan *False Positive* (FP) sebanyak 163 Record. Berdasarkan data pada tabel 2, maka dapat dihitung nilai *accuracy*, *sensitivity*, *specifity*, *ppv* dan *npv*

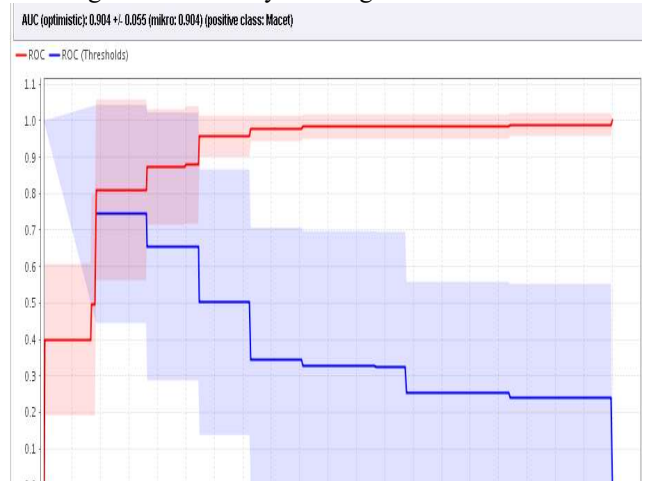
Tabel 2. Hasil Nilai acc, sensitivity, specifity, ppv dan npv C4.5

	Hasil
<i>Accuracy</i>	0,836
<i>Sensitivity</i>	0,778
<i>Specificity</i>	0,871
<i>PPV</i>	0,785
<i>NPV</i>	0,867

Sumber: Hasil Penelitian (2015)

## 3) Evaluasi dengan ROC

Pada gambar 2 menunjukkan grafik ROC dengan nilai AUC (*Area Under Curve*) sebesar 0.904 sehingga Akurasi disebut memiliki tingkat klasifikasi yaitu sangat baik.



Sumber: Hasil Penelitian (2015)

Gambar 2. *Area Under Curve*(AUC)

#### 4) Evaluasi dan Validasi Hasil

Hasil dari pengujian diatas dengan menggunakan *confusion matrix* maupun *ROC curve* terbukti bahwa hasil pengujian algoritma seleksi atribut C4.5 sangat baik. Untuk klasifikasi *data mining*, nilai *Area Under Curve (AUC)* dapat dibagi menjadi beberapa kelompok [2].

- A. 0.90-1.00 = klasifikasi sangat baik
- B. 0.80-0.90 = klasifikasi baik
- C. 0.70-0.80 = klasifikasi cukup
- D. 0.60-0.70 = klasifikasi buruk
- E. 0.50-0.60 = klasifikasi salah

Nilai *Area Under Curve (AUC)* untuk model algoritma C4.5 dengan seleksi atribut sebesar 0.904 dan nilai *accuracy* sebesar 83.67% dengan klasifikasi baik.

#### V. KESIMPULAN

Kesimpulan yang diperoleh setelah melakukan penelitian kelayakan kredit

1. Hasil penelitian untuk nilai akurasi algoritma klasifikasi C4.5 adalah 83.67%.
2. Hasil evaluasi menggunakan *ROC Curve* nilai AUC adalah 0.904 dengan tingkat klasifikasi sangat baik

Berdasarkan proses pengujian dan kesimpulan yang telah dilakukan, maka ada beberapa saran dalam penelitian ini yaitu:

Aspek Manajerial.

- a. Jumlah data yang digunakan hanya 300 dengan jumlah atribut 9, sehingga untuk hasil pengukuran yang lebih baik lagi disarankan untuk menambah jumlah data yang lebih besar dan jumlah atribut yang lebih banyak.
- b. Gunakan metode optimasi lain seperti, *Genetic Algorithm (GA)*, *Ant Colony Optimization* dan lainnya.

Aspek Sistem.

- a. Meningkatkan sistem analisa kredit untuk penentuan kelayakan pemberian kredit bagi calon debitur yang akan mengajukan kredit pada Bank Mayapada.
- b. Mengukur apakah model yang telah dikembangkan berhasil atau tidak, evaluasi digunakan untuk mengukur keakuratan hasil yang dicapai oleh model.

#### REFERENSI

- [1] Anbarasi, M., Anupriya, E., & Iyengar, N. Enchaced 5370-Prediction of Heart Disease with Feature Subset Selection using Genetic Algorithm. *International Journal of engineering Science and Technology* , 5370-5376. 2010.
- [2] Gorunescu, Florin. *Data Mining: Concepts, Models, and Techniques*. Verlag Berlin Heidelberg: Springer . 2011.
- [3] Han, J., & Kamber, M. *Data Mining Concept and Tehniques*. San Fransisco: Morgan Kauffman. 2006.
- [4] Larose Larose, D. T. *Discovering Knowledge in Data*. New Jersey: John Willey & Sons,Inc. 2005.
- [5] Witten, I. H., Frank, E., & Hall, M. A. *Data Mining Practical Machine Learning Tools and Techniques*. USA: Morgan Kaufmann Publishers. 2007.
- [6] Wu, X., & Kumar, V. *The Top Ten Algorithms in Data Mining*. New York: Chapman & Hall/CRC. 2009.

- [7] Vercellis, Carlo. *Business Intelligent:Data Mining and Optimization forDecision Making*. Southern Gate,Chichester, West Sussex: John Willey &Sons, Ltd. 2009.
- [8] Zurada, Could Decision Trees Improve the Classification Accuracy and Interpretability of Loan Granting Decisions. HICSS '10 Proceedings of the 2010 43rd Hawaii International Conference on System Sciences, (hal. 19). Koloa 2010



Nandang Iriadi, M.Kom. Tahun 2002 lulus dari Program Strata Satu (S1) Program Studi Teknik Informatika Universitas Respati Indonesia Jakarta. Tahun 2002 lulus dari Program Strata Dua (S2) Program Studi Magister Ilmu Komputer STMIK Nusa Mandiri Jakarta. Tahun 2010 sudah tersertifikasi dosen dengan Jabatan Fungsional Akademik Lektor di AMIK BSI Jakarta. Aktif mengikuti seminar dan menulis paper di beberapa jurnal diantaranya Jurnal PARADIGMA AMIK BSI Jakarta. Aktif sebagai anggota Forum Akademisi Indonesia.



Nia Nuraeni, M.Kom. Tahun 2012 lulus dari Program Strata Satu (S1) Program Studi Sistem Informasi STMIK Nusa Mandiri Jakarta. Lulus dari Program Strata Dua (S2) Program Studi Magister Ilmu Komputer STMIK Nusa Mandiri Jakarta Tahun 2015. Saat ini bekerja sebagai tenaga pengajar di AMIK BSI Jakarta.

