

EVALUASI PENENTUAN KELAYAKAN PEMBERIAN KREDIT KOPERASI SYARIAH MENGGUNAKAN ALGORITMA KLASIFIKASI C4.5

Siti Masripah

Manajemen Informatika
AMIK BSI Jakarta
Jl. R.S. Fatmawati No.24 Pondok Labu, Jakarta Selatan.
E-mail: siti.stm@bsi.ac.id

Abstract — *Credit is the provision of money or bills that can be equivalent, based on the approval of borrowing between banks and other parties that requires the borrower to pay off their debts after a certain period of time with interest. While the cooperative is a business entity consisting of one person or legal entity with bases its activities based on the principle of cooperation as well as driving people's economy based on the principle of kinship. And to Cooperative Financial Services, here in after referred Syariah is a cooperative whose main business is engaged in financing, investments, and deposits according to the pattern of results (sharia). In analyzing a credit analyst sometimes an inaccurate analysis, so there are some customers who are less capable in performing loan payments, resulting in bad credit. Many studies using C4.5 algorithm for the determination of credit worthiness, but the resulting accuracy is less accurate. After testing the two models namely algorithm C4.5 the results obtained are the C4.5 algorithm produces an accuracy value by 88% and the AUC value of 0.898 with the diagnosis of Good Classification.*

Intisari — Kredit adalah penyediaan uang atau tagihan yang dapat dipersamakan dengan itu, berdasarkan persetujuan pinjam meminjam antara bank dengan pihak lain yang mewajibkan pihak peminjam melunasi hutangnya setelah jangka waktu tertentu dengan pemberian bunga. Sedangkan koperasi adalah badan usaha yang beranggotakan orang-seorang atau badan hukum dengan melandaskan kegiatannya berdasarkan prinsip koperasi sekaligus sebagai penggerak ekonomi rakyat yang berdasarkan asas kekeluargaan. Dan untuk Koperasi Jasa Keuangan Syariah adalah koperasi yang kegiatan usahanya bergerak dibidang pembiayaan, investasi, dan simpanan sesuai pola bagi hasil (syariah). Dalam menganalisa sebuah kredit terkadang seorang analis melakukan analisa tidak akurat, sehingga ada beberapa nasabah yang kurang mampu dalam melakukan pembayaran kredit, sehingga mengakibatkan kredit macet. Setelah dilakukan pengujian dengan model yaitu Algoritma C4.5 maka hasil yang didapat adalah Algoritma C4.5

menghasilkan nilai akurasi sebesar 88% dan nilai AUC sebesar 0,898 dengan tingkat diagnosa *Good Classification*.

Kata Kunci : Algoritma C4.5, Kredit, Koperasi Syariah

PENDAHULUAN

Kredit adalah penyediaan uang atau tagihan yang dapat dipersamakan dengan itu, berdasarkan persetujuan pinjam meminjam antara bank dengan pihak lain yang mewajibkan pihak peminjam melunasi hutangnya setelah jangka waktu tertentu dengan pemberian bunga (UU Perbankan No.10 Tahun 1998). Koperasi adalah badan usaha yang beranggotakan orang-seorang atau badan hukum dengan melandaskan kegiatannya berdasarkan prinsip koperasi sekaligus sebagai penggerak ekonomi rakyat yang berdasarkan asas kekeluargaan. Sedangkan untuk Koperasi Jasa Keuangan Syariah selanjutnya disebut KJKS adalah koperasi yang kegiatan usahanya bergerak dibidang pembiayaan, investasi, dan simpanan sesuai pola bagi hasil (syariah) (KepMen No 91, 2004).

Pembiayaan adalah peyediaan dana oleh pemerintah, dunia usaha, dan masyarakat melalui lembaga keuangan bank, lembaga keuangan bukan bank, atau melalui lembaga lain dalam rangka memperkuat pemodaln usaha kecil (UU RI nomor 9 tahun 1995 pasal 1). Seperti halnya bank, koperasi syariah sebagai pemberi dana dalam melakukan proses pemberian pembiayaan secara garis besar yaitu pengajuan pembiayaan, analisis usulan pembiayaan, persetujuan komite pembiayaan koperasi syariah, pengikatan pembiayaan, dan pencairan dana.

Setiap peminjam atau debitor harus melakukan proses diatas, salah satu yang berperan adalah proses analisis usulan pembiayaan yang dilakukan oleh yang berwenang karena dengan melakukan analisis inilah pihak pembiayaan akan melakukan pencairan dana, jika peminjam memiliki nilai baik setelah dilakukan analisa, sehingga akan

mengurangi resiko kredit yang akan diterima oleh pemberi dana. Dalam memberikan kredit atau pinjaman saat proses mengembalikan dana tidak selalu berjalan lancar, pasti akan mengalami resiko kredit yaitu ketidak sanggupannya nasabah dalam membayar angsuran. Banyak faktor yang menentukan nasabah tidak lancar dalam membayar angsuran kredit salah satu contoh adalah situasi ekonomi yang terjadi sehingga mengganggu usaha yang dijalankannya. Dalam arti luas resiko kredit adalah ketidak pastian atau fluktuasi laba dalam kegiatan kredit (Yu, Chen, Koronios, Zhu, & Guo, 2007). Untuk mengurangi resiko kredit maka analisa kredit merupakan hal yang penting dalam manajemen resiko keuangan (Lai, Yu, Zhou, & Wang, 2006).

Data histori merupakan data latihan atau data pengalaman, karena dengan data tersebut kita akan berlatih untuk mendapatkan pengetahuan. Algoritma klasifikasi akan menggunakan data latihan sehingga akan menghasilkan pengetahuan, pengetahuan untuk menggolongkan risiko kredit seorang nasabah pada masa mendatang berdasarkan variabel-variabel yang ada. Sebagai bahan tolak ukur nasabah disetujui atau ditolak, dapat melihat dari data histori kredit nasabah pada koperasi syariah. Dibawah ini merupakan grafik yang menunjukkan bahwa nasabah yang bermasalah dalam hal pembayaran angsuran kredit lebih besar dibandingkan dengan nasabah yang lancar dalam pembayaran angsuran kredit, data diambil berdasarkan tahun 2010.

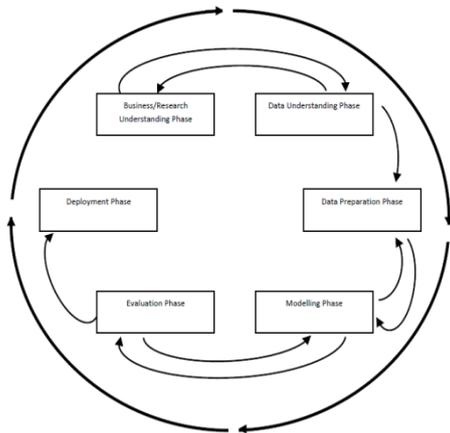
BAHAN DAN METODE

Dalam penelitian ini, digunakan beberapa media diantaranya buku, jurnal baik jurnal nasional maupun internasional serta prosiding sebagai referensi untuk menjelaskan tentang perkreditan, data mining, model algoritma C4.5, metode evaluasi *confusion matrix* dan *ROC curve*. Pada pasal 1 angka 11 Undang-undang no. 10 tahun 1998 tentang perubahan undang-undang no 7 tahun 1992 tentang perbankan yang dimaksud dengan kredit adalah penyediaan uang atau tagihan yang dapat dipersamakan dengan itu, berdasarkan persetujuan pinjam meminjam antara bank dengan pihak lain yang mewajibkan pihak peminjam melunasi hutangnya setelah jangka waktu tertentu dengan pemberian bunga.

Seperti halnya bank, Kredit Jasa Keuangan Syariah sebagai pemberi dana dalam melakukan penilaian kredit terhadap nasabah, ada beberapa kriteria yang harus dilakukan atau prinsip utama yang berkaitan dengan kondisi calon nasabah. Prinsip ini dikenal dengan 5C yaitu:

1. *Character*
Penilaian terhadap karakter atau kepribadian calon peminjam untuk memperkirakan kemungkinan bahwa peminjam dapat memenuhi kewajibannya.
2. *Capacity*
Penilaian tentang kemampuan peminjam untuk melakukan pembayaran. Kemampuan diukur dengan catatan prestasi peminjam dimasa lalu yang didukung dengan mengamati dilapangan atas sarana dan usahanya seperti karyawan, mesin, sarana produksi, cara usaha, dan lain sebagainya.
3. *Capital*
Penilaian terhadap kemampuan modal yang dimiliki oleh calon peminjam, diukur dengan posisi usaha/perusahaan secara keseluruhan yang ditunjukkan oleh rasio keuangan dan penekanan pada komposisi modal.
4. *Colateral*
Jaminan yang dimiliki calon peminjam. Penilaian ini untuk lebih meyakinkan bahwa jika suatu risiko kegagalan pembayaran tercapai terjadi, maka jaminan dapat dipakai sebagai pengganti kewajiban.
5. *Conditions*
Pihak koperasi syariah harus melihat kondisi ekonomi yang terjadi dimasyarakat dan secara spesifik melihat adanya keterkaitan dengan jenis usaha yang dilakukan oleh calon peminjam. Hal tersebut dilakukan karena kondisi eksternal memiliki pengaruh yang cukup besar dalam proses berjalannya usaha calon peminjam dalam jangka panjang.

Data mining adalah proses menemukan korelasi baru yang bermakna, pola dan tren dengan memilah-milah sejumlah besar data yang tersimpan dalam repositoru, menggunakan teknologi penalaran pola serta teknik-teknik statistik dan matematika (Larose, 2005). Data mining adalah sebuah proses, sehingga dalam melakukan prosesnya harus sesuai prosedur yaitu proses CRISP-DM (*Cross-Industry Standard Process for Data Mining*) yaitu sebagai keseluruhan proses, preproccesing data, pembentukan model, model evaluasi, dan akhirnya penyebaran model (Larose, 2005). Dapat dilihat pada gambar 1 sebagai berikut:



Sumber : Gorunescu (2011)

Gambar 1. Proses CRISP-DM

Pada Gambar 1. CRISP-DM memiliki enam proses atau phase yaitu:

1. *Business understanding*
Pada tahap pertama bisa disebut juga tahap pemahaman penelitian, menentukan tujuan proyek penelitian dalam perumusan mendefinisikan masalah data mining.
2. *Data Understanding*
Dilakukan pengumpulan data, kemudian menganalisa data serta evaluasi kualitas data.
3. *Data Preparation*
Persiapkan data mentah kemudian di seting untuk data akhir yang akan digunakan untuk fase selanjutnya, pilih kasus dan variabel yang diinginkan yang digunakan untuk menganalisa sesuai analisa masalah, lakukan transformasi pada variabel tertentu jika diinginkan, bersihkan data untuk alat pemodelan
4. *Modelling*
Pada tahap ini, pilih dan terapkan teknik pemodelan yang tepat, lakukan pengaturan model untuk mengoptimalkan hasil, jika diperlukan lakukan ulang ke tahap persiapan sesuai dengan persyaratan spesifikasi dari teknik data mining tertentu
5. *Evaluasi*
Melakukan evaluasi satu atau lebih model, tentukan apakah model sudah mencapai tujuan yang diterapkan dalam tahap pertama, mengambil keputusan mengenai penggunaan hasil data mining
6. *Deployment*
Memanfaatkan model yang telah dibuat, *deployment* yang sederhana adalah sampai menghasilkan laporan sedangkan *deployment* yang kompleks adalah melaksanakan model untuk proses data mining paralel pada departemen lain.

Fungsi-fungsi dalam data mining mengacu pada Larose (2005) terdapat enam fungsi yaitu (Susanto & Suryadi, 2010):

1. Fungsi deskripsi (*description*)
Fungsi deskripsi adalah cara yang digunakan untuk menggambarkan sekumpulan data secara ringkas. Banyak cara yang digunakan dalam memberikan gambaran secara ringkas bagi sekumpulan data yang besar jumlahnya dan banyak macamnya yaitu Deskripsi Grafis, Deskripsi Lokasi, dan Deskripsi Keragaman
2. Fungsi estimasi (*estimation*)
Fungsi estimasi adalah fungsi untuk memperkirakan suatu hal yang sudah ada datanya. Fungsi estimasi terdiri dari dua cara yaitu Estimasi Titik dan Estimasi Selang Kepercayaan.
3. Fungsi prediksi (*prediction*)
Fungsi prediksi adalah memperkirakan hasil dari hal yang belum diketahui, untuk mendapatkan hal baru yang akan muncul selanjutnya. Cara memprediksi dalam fungsi ini adalah Regresi Linier.
4. Fungsi klasifikasi (*classification*)
Fungsi klasifikasi atau menggolongkan suatu data. Cara yang digunakan terdiri dari algoritma Mean Vector, algoritma K-nearest Neighbor, algoritma ID3, algoritma C4.5, dan algoritma C5.0
5. Fungsi pengelompokan (*cluster*)
Fungsi pengelompokan, data yang dikelompokkan disebut objek atau catatan yang memiliki kemiripan atribut kemudian dikelompokkan pada kelompok yang berbeda. Algoritma yang digunakan adalah algoritma Hirarchical Clustering, algoritma Partitional Clustering, algoritma Single Linkage, algoritma Complete Linkage, algoritma Average Linkage, algoritma K-Means dan lain-lain.
6. Fungsi asosiasi (*association*)
Fungsi asosiasi adalah untuk menemukan aturan asosiasi (*association rule*) yang mampu mengidentifikasi item-item yang menjadi objek. Algoritma yang digunakan adalah algoritma Generalized Association Rules, Quantitative Association Rule, asynchronous Parallel Mining

Salah satu metode klasifikasi menarik yang melibatkan konstruksi pohon keputusan, koleksi node keputusan, terhubung oleh cabang-cabang, memperpanjang bawah dari simpul akar samapai berakhir di node daun. Dimulai dari node root, yang oleh konvensi ditempatkan dibagian atas dari diagram pohon keputusan, atribut diuji pada node keputusan, dengan setiap hasil yang mungkin menghasilkan cabang. Setiap cabang

kemudian mengarah ke node lain baik keputusan atau ke node daun untuk mengakhiri (Larose, 2005).

Algoritma C4.5 merupakan bagian dari kelompok algoritma *decision trees* dan merupakan katerogi 10 algoritma yang paling populer. Diakhir tahun 1970 hingga awal tahun 1980-an, J.Rosss Quinlan seorang peneliti dibidang mesin pembelajaran mengembangkan sebuah model keputusan yang dinamakan ID3 (*Iterative Dichotomiser*), walaupun sebelumnya proyek ini telah dibuat oleh E.B. Hunt, J. Marin, dan P.T. Stone. Quinlan kemudian membuat algoritma C4.5 (pengembangan dari ID3) yang berbasis *supervised learning* (Han & Kamber, 2006).

Untuk melakukan evaluasi terhadap model klasifikasi berdasarkan perhitungan objek testing mana yang diprediksi benar dan tidak benar. Perhitungan ini ditabulasikan kedalam tabel yang disebut *confusion matrix* (Gorunescu, 2011). Bentuk *confusion matrix* dapat dilihat pada Tabel 1 berikut ini:

Tabel 1. Confusion Matrix

CLASSIFICATION	PREDICTED CLASS	
	Class = YES	Class = NO
OBSERVED	Class = YES	Class = NO
CLASS	(true positive-TP)	(false negative-FN)
	Class = No	(false positive-FP)

Sumber : Gorunescu (2011)

Pada Tabel 1, untuk *True positive* merupakan tupel positif didata set yang diklasifikasikan positif, *True negatives* merupakan tupel negatif di data set yang diklasifikasikan negatif. *False positives* adalah tupel positif di data set yang diklasifikasikan negatif *False negatives* merupakan jumlah tupel negatif yang diklasifikasikan positif.

Setelah dilakukan *confusion matrix* berikutnya akan dihitung *accuracy*, *sensitivity*, *specificity*, *PPV*, *NPV*. *Sensitivity* digunakan untuk membandingkan jumlah *true positives* terhadap jumlah tupel yang *positives* sedangkan *specificity* adalah perbandingan jumlah *true negatives* terhadap jumlah tupel yang *negatives*. Sedangkan untuk *PPV*(nilai prediktif positif) adalah proporsi kasus dengan hasil diagnosa positif, *NPV*(nilai prediktif negatif) adalah proporsi kasus dengan hasil diagnosa negatif. Berikut perhitungannya:

$$Accuracy = \frac{a+d}{a+b+c+d} = \frac{TP+TN}{TP+TN+FP+FN}$$

$$Sensitivity = \frac{\text{number of TruePositives}}{\text{Number of TruePositives} + \text{Number of FalseNegatives}}$$

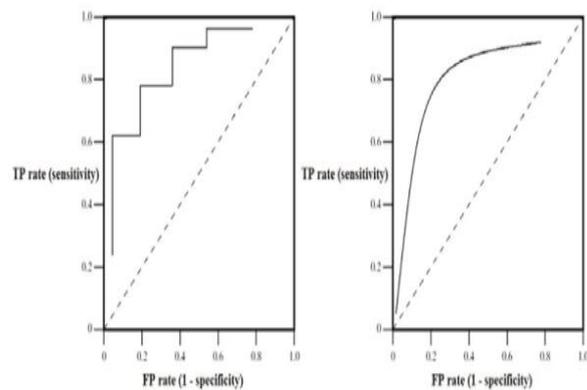
$$Specificity = \frac{\text{number of TrueNegatives}}{\text{Number of TrueNegatives} + \text{Number of FalsePositives}}$$

$$PPV = \frac{\text{number of TruePositives}}{\text{Number of TruePositives} + \text{Number of FalsePositives}}$$

$$NPV = \frac{\text{number of TrueNegatives}}{\text{Number of TrueNegatives} + \text{Number of FalseNegatives}}$$

Sensitivity juga dapat dikatakan *true positive rate* (TP rate) atau *recall*. Sebuah *sensitivity* 100% berarti bahwa pengklasifikasian mengakui sebuah kasus yang diamati positif. Misalnya semua orang memiliki kanker ganas diakui sebagai sakit.

Kurva ROC (*Receiver Operating Characteristic*) adalah ilustrasi grafis dari kemampuan diskriminan dan biasanya diterapkan untuk masalah klasifikasi biner (Yu, Chen, Koronios, Zhu, & Guo, 2007). Secara teknik, kurva ROC juga disebut grafik ROC, dua dimensi grafik yaitu TP rate diletakan pada sumbu Y, sedangkan FP rate diletakan pada sumbu X. Grafik ROC menggambarkan trade-off antara manfaat (*true positives*) dan biaya (*false positives*). Berikut tampilan dua jenis kurva ROC (*discrete* dan *continuous*).



Sumber : Gorunescu (2011)

Gambar 2. Grafik ROC (Discreate dan Continuous)

Dari Gambar 2 ada beberapa yang penting untuk dicatat. Titik kiri bawah (0,0) yaitu diantara nilai TP dan FP, titik (1,1) merupakan klasifikasi positif. Titik (0,1) merupakan klasifikasi sempurna (yaitu tidak ada FN dan tidak ada FP). yang benar-benar acak akan memberikan titik sepanjang garis diagonal dari kiri bawah kesudut kanan atas. Garis ini membagi ruang ROC sebagai berikut:

- (a) poin diatas garis diagonal merupakan hasil klasifikasi yang baik
- (b) poin dibawah garis diagonal merupakan hasil klasifikasi yang buruk

dapat disimpulkan satu poin untuk ROC adalah lebih baik dari pada yang lainnya jika arah garis melintang dari kiri bawah ke kanan atas didalam grafik tersebut.

Metode Penelitian

Pada umumnya ada empat metode penelitian yang digunakan yaitu *Action Reserch*, *Experiment*, *Case Study*, dan *Survey* (Dawson, 2009). Adapun metode penelitian yang digunakan adalah bentuk penelitian *Experiment*. Penelitian eksperimen merupakan sebuah penyelidikan hubungan kausal menggunakan tes dikendalikan oleh peneliti (Dawson, 2009).

Dalam eksperimen biasanya terdiri dari:

1. Mendefinisikan hipotesis teoritis
2. Memilih sampel dari populasi yang diketahui
3. Mengalokasikan sampel untuk kondisi percobaan yang berbeda
4. Memperkenalkan perubahan yang direncanakan untuk satu atau lebih variabel
5. Mengukur sejumlah kecil variabel
6. Mengontrol semua variabel

Penelitian eksperimen biasanya dilakukan dalam proyek pengembangan, evaluasi dan pemecahan masalah (Dawson, 2009). Dalam penelitian eksperimen digunakan spesifikasi hardware dan software sebagai alat bantu dalam penelitian yaitu terdapat pada Tabel 2:

Tabel 2. Spesifikasi Hardware dan Software

Hardware	Software
CPU : Atom Dual Core	Sistem Operasi : Windows 7
Memory : 2 GB	Data Mining : Rapid Miner 5.2
Hardisk : 120 GB	

Sumber : penulis (2014)

Dalam metode penelitian eksperimen, digunakan model proses CRISP-DM (*Cross-Standard Industry Process for Data Mining*) yang terdiri dari 6 tahapan (Larose, 2005):

1. *Bussiness understanding*
2. *Data understanding*
3. *Data preparation*
4. *Modelling*
5. *Evaluation*
6. *Deployment*

HASIL DAN PEMBAHASAN

Dalam penelitian ini populasi sebesar 866 dari 866 data dibagi menjadi dua bagian, 766 digunakan untuk data training dan 100 digunakan untuk data testing. Data sebanyak 866 dan atribut yang terdiri dari 44 atribut, akan dilakukan beberapa penyeleksian untuk menghasilkan data yang dibutuhkan, tahapannya yaitu:

1. *Data Cleaning* untuk membersihkan nilai yang kosong atau tupel yang kosong. Sebagai contoh atribut tunggakan denda.
2. *Data Integration* yang berfungsi menyatukan tempat penyimpanan yang berbeda kedalam satu data. Dalam kasus ini hanya ada satu tempat penyimpanan data yaitu status kredit nasabah.
3. *Data reduction* jumlah atribut yang digunakan mungkin terlalu besar, dari 44 atribut yang digunakan hanya 17 atribut yang diperlukan, dan atribut yang tidak diperlukan akan dihapus.

Berdasarkan data tersebut maka didapat candidate split yaitu tunggakan pokok, jumlah pinjaman, jumlah angsuran perbulan, tunggakan bunga, saldo nominatif, Sehingga dapat digambarkan rule nilainya pada Tabel 3 sebagai berikut:

Tabel 3. Candidate split dan rule nilai atribut algoritma C4.5

Candidate split	Child nodes
1 Tunggakan pokok ≤ 7055.330 ≤ 9000 ≤ 166833.330 ≤ 313750.005 ≤ 606600	Tunggakan pokok >7055.330 > 9000 > 166833.330 > 313750.005 > 606600
2 Jml pinjaman ≤ 1831667 ≤ 505626 ≤ 925208.380 ≤ 14692499.500	Jml pinjaman > 1831667 > 505626 > 925208.380 > 14692499.500
3 Jml angsuran per bulan ≤ 12750 ≤ 239633.205	Jml angsuran per bulan > 12750 > 239633.205

4	Tunggakan bunga ≤ 1756 ≤ 9000 ≤ 15000	Tunggakan bunga > 1756 > 9000 > 15000
	≤ 41250 ≤ 112500	> 41250 > 112500
5	Saldo nominatif ≤ 433750 ≤ 950000 ≤ 1669500.010	Saldo nominatif > 433750 > 950000 > 1669500.010
6	Jkw ≤ 7.500 ≤ 22.500	Jkw > 7.500 > 22.500
7	Bi golongan penjamin = 000 Bi golongan penjamin = 875	

Sumber : Data Olahan (2014)

Tahap selanjutnya dilakukan pemrosesan data training sehingga akan menghasilkan beberapa aturan dan akan membentuk sebuah pohon keputusan. Berikut langkah-langkah yang akan dilakukan.

1. Menghitung jumlah kasus class LANCAR dan class MACET dan *Entropy* dari semua kasus dan kasus yang dibagi berdasarkan atribut di Tabel 3.3. Baris total *entropy* dihitung berdasarkan data training menggunakan Persamaan sebagai berikut:

$$\text{Entropy(Total)} = \left(-\frac{210}{766} \cdot \log_2\left(\frac{210}{766}\right)\right) + \left(-\frac{556}{766} \cdot \log_2\left(\frac{556}{766}\right)\right)$$

$$\text{Entropy(Total)} = 0,8473$$

Sumber : Data Olahan (2014)

2. Kemudian hitung masing-masing Gain berdasarkan atribut pada Tabel 3.3 diatas, sebagai contoh untuk tunggakan pokok. Untuk record tunggakan pokok ≤ 7055.330 terdiri dari 143 LANCAR dan 1 MACET, untuk tunggakan pokok > 7055.330 terdiri dari 67 LANCAR dan 555 MACET. Dapat dihitung Entropinya sebagai berikut:

$$H_{\text{Tunggakan Pokok} \leq 7055.330} = \left(-\frac{143}{144} \log_2\left(\frac{143}{144}\right)\right) + \left(-\frac{1}{144} \log_2\left(\frac{1}{144}\right)\right)$$

$$H_{\text{Tunggakan Pokok} > 7055.330} = \left(-\frac{67}{622} \log_2\left(\frac{67}{622}\right)\right) + \left(-\frac{555}{622} \log_2\left(\frac{555}{622}\right)\right)$$

$$H_{\text{Tunggakan Pokok} \leq 7055.330(T)} = \frac{144}{766} \cdot 0.0598 + \frac{622}{766} \cdot 0.493$$

$$\text{Entropy(Tunggakan pokok)} = 0.8473 - 0.4115 = 0.4358$$

Sumber : Data Olahan (2014)

Untuk data selengkapnya dapat dilihat Tabel dibawah ini

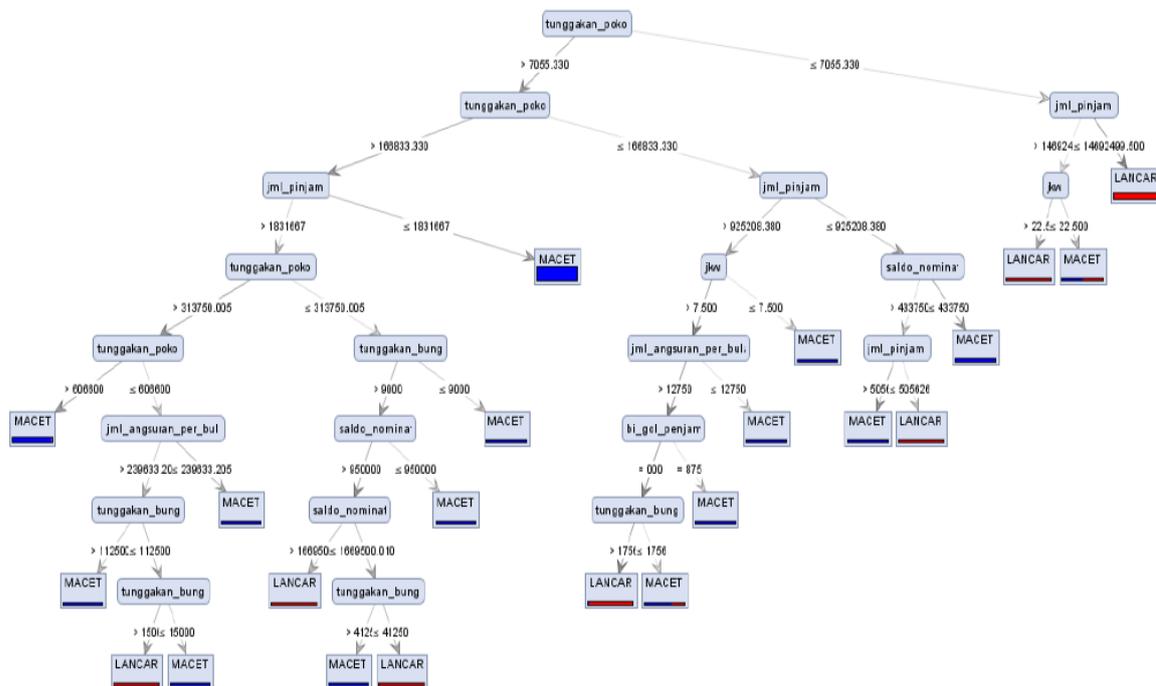
Tabel 4. Informasi Gain

Kandidat Split	Child Nodes	Informasi Gain (Entropy Reduction)
1	Tunggakan pokok ≤ 7055.330 dan > 7055.330	0.4358
2	Tunggakan pokok ≤ 166833.330 dan > 166833.330	0.4157
3	Tunggakan pokok ≤ 313750.005 dan > 313750.005	0.3039
4	Tunggakan pokok ≤ 606600 dan > 606600	0.15
5	Jumlah pinjaman ≤ 505626 dan > 505626	0.0256
6	Jumlah pinjaman ≤ 925208.380 dan > 925208.380	0.0876
7	Jumlah pinjaman ≤ 1831667 dan > 1831667	0.2908
8	Jumlah pinjaman ≤ 14692499.500 dan > 14692499.500	0.0014
9	Jumlah angsuran ≤ 12750 dan > 12750	0.0012
10	Jumlah angsuran ≤ 239633.205 dan > 239633.205	-0.0869

11	Tunggakan bunga ≤ 1756 dan > 1756	0.0302
12	Tunggakan bunga ≤ 9000 dan > 9000	-0.1917
13	Tunggakan bunga ≤ 15000 dan > 15000	0.0.2193
14	Tunggakan bunga ≤ 41250 dan > 41250	0.1124
15	Tunggakan bunga ≤ 112500 dan > 112500	0.0524
16	Saldo nominatif ≤ 433750 dan > 433750	0.0038
17	Saldo nominatif ≤ 950000 dan > 950000	0.0176
18	Saldo nominatif ≤ 1669500.010 dan > 1669500.010	0.133
19	Jkw ≤ 7.500 dan > 7.500	0
20	Jkw ≤ 22.500 dan > 22.500	0.0011
21	Bi gol penjamin = 000	0.1977
22	Bi gol penjamin = 875	0.7533

Sumber : Data Olahan (2014)

Berdasarkan tabel diatas, maka terbentuklah digram pohon keputusan sebagai berikut:



Sumber : Data Olahan (2014)

Gambar 3. Pohon Keputusan Klasifikasi Nasabah

Pada Gambar 3 dapat dijelaskan sebagai berikut: pada root tunggakan pokok jika bernilai ≤ 7055.330 dan jml pinjaman ≤ 14692499.500 maka menghasilkan daun atau class = lancar. Untuk rincian lebih lengkap dari pohon keputusan pada Gambar 3.1 diatas dapat dilihat aturan sebagai berikut:

1. R1: IF tunggakan pokok ≤ 7055.330 AND jml pinjaman ≤ 14692499.500 THEN class = LANCAR
2. R2: IF tunggakan pokok ≤ 7055.330 AND jml pinjaman > 14692499.500 AND jkw ≤ 22.500 THEN class = MACET
3. R3: IF tunggakan pokok ≤ 7055.330 AND jml pinjaman > 14692499.500 AND jkw > 22.500 THEN class = LANCAR
4. R4: IF tunggakan pokok > 7055.330 AND tunggakan pokok ≤ 166833.330 AND jml pinjaman ≤ 925208.380 AND saldo nominal ≤ 433750 THEN class = MACET
5. R5: IF tunggakan pokok > 7055.330 AND tunggakan pokok ≤ 166833.330 AND jml pinjaman ≤ 925208.380 AND saldo nominal > 433750 AND jumlah pinjaman ≤ 505626 THEN class = LANCAR
6. R6: IF tunggakan pokok > 7055.330 AND tunggakan pokok ≤ 166833.330 AND jml pinjaman ≤ 925208.380 AND saldo nominal > 433750 AND jumlah pinjaman > 505626 THEN class = MACET
7. R7: IF tunggakan pokok > 7055.330 AND tunggakan pokok ≤ 166833.330 AND jml pinjaman > 925208.380 AND jkw ≤ 7.500 THEN class = MACET
8. R8: IF tunggakan pokok > 7055.330 AND tunggakan pokok ≤ 166833.330 AND jml pinjaman > 925208.380 AND jkw > 7.500 AND jml angsuran perbulan ≤ 12750 THEN class = MACET
9. R9: IF tunggakan pokok > 7055.330 AND tunggakan pokok ≤ 166833.330 AND jml pinjaman > 925208.380 AND jkw > 7.500 AND jml angsuran perbulan > 12750 AND bi gol penjamin = 875 THEN class = MACET
10. R10: IF tunggakan pokok > 7055.330 AND tunggakan pokok ≤ 166833.330 AND jml pinjaman > 925208.380 AND jkw > 7.500 AND jml angsuran perbulan > 12750 AND bi gol penjamin = 000 AND tunggakan bunga ≤ 1756 THEN class = MACET
11. R11: IF tunggakan pokok > 7055.330 AND tunggakan pokok ≤ 166833.330 AND jml pinjaman > 925208.380 AND jkw > 7.500 AND jml angsuran perbulan > 12750 AND bi gol penjamin = 000 AND tunggakan bunga > 1756 THEN class = LANCAR
12. R12: IF tunggakan pokok > 166833.330 AND jml pinjaman ≤ 1831667 THEN class = MACET
13. R13: IF tunggakan pokok > 166833.330 AND jml pinjaman > 1831667 AND tunggakan pokok ≤ 313750.005 AND tunggakan bunga ≤ 9000 THEN class = MACET
14. R14: IF tunggakan pokok > 166833.330 AND jml pinjaman > 1831667 AND tunggakan pokok ≤ 313750.005 AND tunggakan bunga > 9000 AND saldo nominal ≤ 950000 THEN class = MACET
15. R15: IF tunggakan pokok > 166833.330 AND jml pinjaman > 1831667 AND tunggakan pokok ≤ 313750.005 AND tunggakan bunga > 9000 AND saldo nominal > 950000 AND saldo nominal ≤ 1669500.010 AND tunggakan bunga ≤ 41250 THEN class = LANCAR
16. R16: IF tunggakan pokok > 166833.330 AND jml pinjaman > 1831667 AND tunggakan pokok ≤ 313750.005 AND tunggakan bunga > 9000 AND saldo nominal > 950000 AND saldo nominal ≤ 1669500.010 AND tunggakan bunga > 41250 THEN class = MACET
17. R17: IF tunggakan pokok > 166833.330 AND jml pinjaman > 1831667 AND tunggakan pokok ≤ 313750.005 AND tunggakan bunga > 9000 AND saldo nominal > 950000 AND saldo nominal > 1669500.010 THEN class = LANCAR
18. R18: IF tunggakan pokok > 313750.005 AND tunggakan pokok > 606600 THEN class = MACET
19. R19: IF tunggakan pokok > 313750.005 AND tunggakan pokok ≤ 606600 AND jml angsuran per bulan ≤ 239633.205 THEN class = MACET
20. R20: IF tunggakan pokok > 313750.005 AND tunggakan pokok ≤ 606600 AND jml angsuran per bulan > 239633.205 AND tunggakan bunga > 112500 THEN class = MACET
21. R21: IF tunggakan pokok > 313750.005 AND tunggakan pokok ≤ 606600 AND jml angsuran per bulan > 239633.205 AND tunggakan bunga ≤ 112500 AND tunggakan bunga > 15000 THEN class = LANCAR
22. R22: IF tunggakan pokok > 313750.005 AND tunggakan pokok ≤ 606600 AND jml angsuran per bulan > 239633.205 AND tunggakan

bunga ≤ 112500 AND tunggakan bunga ≤ 15000 THEN class = MACET

Hasil dari pengujian model yang dilakukan adalah penentuan kelayakan kredit dengan algoritma klasifikasi C4.5. Hasil dari uji coba yang dilakukan yaitu untuk menghasilkan nilai *accuracy* dan nilai AUC (*Area Under Curve*).

a. 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, kemudian masukan data testing yang sudah disiapkan kedalam *confusion matrix* sehingga didapatkan hasil pada Tabel 4 dibawah ini:

Tabel 4. Konversi confusion matrix algoritma klasifikasi C4.5

Observed Class	Predicate Class	
	YES	NO
YES	50	3
NO	9	38

Sumber : Data Olahan (2014)

Berdasarkan data yang terdapat pada *confusion matrix* diatas maka dapat kita hitung untuk mencari nilai *accuracy*, *sensitivity*, *specifity*, *ppv*, dan *npv*, hasil dapat dilihat pada Tabel 5 dibawah ini:

Tabel 5. Nilai sensitivity, specificity, ppv, npv, dan accuracy

	Nilai (%)
<i>Accuracy</i>	88
<i>Sensitivity</i>	94,34
<i>Specificity</i>	80,85
<i>PPV</i>	84,75
<i>NPV</i>	92,68

Berdasarkan Tabel 5 menunjukan bahwa, tingkat akurasi menggunakan algoritma klasifikasi C4.5 adalah sebesar 88%.

b. Evaluasi Dengan ROC Curve

Pada Gambar 4 menunjukan grafik ROC dengan nilai AUC (*Area Under Curve*) sebesar 0.898. Akurasi memiliki tingkat diagnosa yaitu (Gorunescu, 2011):

Akurasi bernilai 0.90 - 1.00 = *Excellent classification*

Akurasi bernilai 0.80 - 0.90 = *Good classification*

Akurasi bernilai 0.70 - 0.80 = *Fair classification*

Akurasi bernilai 0.60 - 0.70 = *Poor classification*

Akurasi bernilai 0.50 - 0.60 = *Failure*

Sedangkan hasil yang didapat dari pengolahan ROC yang dapat dilihat pada Gambar 4 sebesar 0.898 dengan tingkat diagnosa *Good classification*.



Sumber : Data Olahan (2014)

Gambar 4 Nilai AUC dalam grafik ROC algoritma C4.5

KESIMPULAN

Berdasarkan hasil penelitian untuk nilai akurasi algoritma klasifikasi C4.5 senilai 88% serta dievaluasi menggunakan ROC curve yaitu, nilai AUC berdasarkan ROC curve untuk algoritma klasifikasi C4.5 bernilai 0.898 dengan tingkat diagnosa *Good classification*,

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

1. Menambahkan jumlah data yang lebih besar dan atribut yang lebih banyak, sehingga hasil pengukuran yang akan didapatkan lebih baik lagi.
2. Menggunakan metode optimasi lain seperti *Ant Colony Optimization (ACO)*, *Genetik Algorithm (GA)*, dan lainnya.
3. Meningkatkan lagi sistem analisa kredit untuk penentuan kelayakan pemberian kredit bagi nasabah yang akan mengajukan kredit pada koperasi syariah.

4. Penerapan kedalam sebuah aplikasi berbasis teknologi informasi guna membantu pihak koperasi dalam melaksanakan analisa kredit.

REFERENSI

- Bastos, J. A. (2008). Credit Scoring with Boosted Decision Trees. *MPRA* , hal 1- 13.
- Gorunescu, F. (2011). *Data Mining Concepts, Model and Techniques*. Berlin: Springer.
- Han, J., & Kamber, M. (2006). *Data Mining Concepts and Techniques*. San Francisco: Diane Cerra.
- Jiang, Y. (2009). Credit Scoring Model Based on Decision Tree and the Simulated Annealing Algorithm. *2009 World Congress on Computer Science and Information Engineering* (hal. 18 - 22). Los Angeles: IEEE Computer Society.
- Larose, D. T. (2005). *Discovering Knowledge In Data*. Canada: Wiley- Interscience.
- Rapid-I. (2010). *Rapid Miner User Manual* . Rapid-I.
- Sugiyono. (2002). *Metode Penelitian Bisnis*. Bandung: CV. Alfabeta.
- Susanto, S., & Suryadi, D. (2010). *Pengantar Data Mining menggali Pengetahuan dari Bongkahan Data*. Yogyakarta: C.V ANDI OFFSET.
- UndangUndang Perbankan No.10 Tahun 1998*.
- Yu, L., Chen, G., Koronios, a., Zhu, S., & Guo, X. (2007). Application and Comparison of Classification Techniques in Controlling Credit Risk. *World Scientific* , 111.
- Zurada, J. (2010). Could decision trees improve the classification accuracy and interpretability of loan granting decision? *43rd Hawaii International Conference on System Sciences* (pp. 1-9).

BIODATA PENULIS



Siti Masripah, M.Kom. Tempat dan tanggal lahir : Bekasi, 27 Desember 1983, Riwayat Pendidikan : S2 Pasca Sarjana STMIK Nusa Mandiri Jakarta, S1 STMIK Nusa Mandiri Jakarta. Pekerjaan : Menjadi Pengajar sejak tahun 2006 di AMIK BSI Jakarta sampai sekarang. Penelitian yang pernah dibuat : Jurnal publikasi tahun 2010 : Tingkat Kelayakan Penerimaan Karyawan Sebuah Perusahaan "XYZ" dengan Metode Fuzzy Logic. Karya ilmiah mandiri yang dipublikasikan pada jurnal PARADIGMA Edisi Volume XII No.2 Bulan September 2010 ISSN No. 1410-5963. Penelitian Tahun 2011 : Penerapan Metode Fuzzy Logic dalam Melakukan Penilaian Kinerja Guru pada Sekolah Dasar Islam Terpadu (SDIT) I'QRO. Dipublikasikan pada jurnal PARADIGMA Edisi Volume XIII No. 2 Bulan September 2011 ISSN No. 1410-5963. Penelitian Tahun 2014 : Application Determination of Credit Feasibility in Sharia Cooperative with C4.5 Algorithm. Dipublikasikan pada Seminar Internasional ISIT 2014 Proceedings Bulan Oktober 2014 ISBN 978-602-99213-8-0.