

KLASIFIKASI CALON NASABAH PEMBIAYAAN PADA PT SINAR MITRA SEPADAN FINANCE MENGGUNAKAN ALGORITMA C4.5

Agung Triayudi¹, Riska Susilawati²

Program Studi Teknik Informatika - Universitas Serang Raya

agung.triayudi@gmail.com riskasusilawati23@gmail.com

Abstrak – Penelitian ini memberikan satu model penentuan kelayakan pemberian fasilitas pembiayaan. Kendala yang sering ditemukan dalam proses pembiayaan di sebuah perusahaan yaitu sulitnya menentukan nasabah mana yang termasuk kedalam kelompok lancar dan tidak lancar. Oleh karena itu diperlukan suatu aplikasi yang membantu perusahaan dalam proses pengklasifikasian nasabah. Untuk menentukan calon nasabah mana yang termasuk kelompok lancar dan tidak lancar. Data dalam penelitian ini diambil dari perusahaan PT. Sinar Mitra Sepadan Finance Cabang Pandeglang dalam jumlah yang cukup banyak proses ini bisa disebut sebagai data mining. Dalam penelitian ini untuk menganalisa data digunakan teknik *decision tree* dengan memanfaatkan algoritma C4.5. Dalam perancangan aplikasi menggunakan diagram konek, DFD, dan penyajian aplikasi menggunakan Borland Delphi 7 serta database Xampp. Hasil yang di capai adalah sebuah Aplikasi Algoritma C4.5 Untuk Pengklasifikasian Calon Nasabah yang Implementasinya dapat membantu bagian Data Enty dalam penyeleksian dan penerimaan calon nasabah.

Kata Kunci : Algoritma C4.5, Nasabah, Data Mining

I. PENDAHULUAN

Dalam beberapa kasus seperti nasabah yang menunggak membuat pendapatan perusahaan menjadi menurun dan tidak stabil, hal ini terjadi karena nasabah tidak membayar tunggakan, Sehingga dengan terpaksa pihak perusahaan melelang agunan yang di ajukan untuk membayar ganti rugi atas kerugian perusahaan. Akan tetapi harga agunan yang dilelang turun drastis sehingga hanya menggantikan sebagian dari kerugian yang dialami oleh perusahaan. Perusahaan menghimbau kepada para kredit marketing untuk berhati – hati dalam memilih calon nasabah, agar kedepannya perusahaan tidak mengalami penurunan modal yang cukup berarti untuk masa depan perusahaan.

Penelitian ini akan difokuskan pada masalah pengklasifikasi calon nasabah pembiayaan. Dengan sistem pengklasifikasi calon nasabah pembiayaan dengan akurat, maka proses selanjutnya yaitu pemilihan calon nasabah lancar akan lebih mudah dengan menggunakan metode algoritma C4.5. Pada penelitian ini juga, peneliti berusaha mengumpulkan data (data mining) nasabah sebuah perusahaan pembiayaan untuk mengetahui lancar atau tidak lancarnya nasabah tersebut. Data yang ada dianalisis menggunakan pendekatan pohon keputusan (decision tree) yaitu algoritma C4.5. Dengan algoritma ini dapat diketahui data nasabah mana yang dikelompokkan ke kelas lancar dan data nasabah mana yang dikelompokkan ke kelas tidak lancar. Kemudian pola tersebut dapat digunakan untuk memperkirakan nasabah yang bergabung, sehingga perusahaan bisa mengambil keputusan menerima atau menolak calon nasabah tersebut.

II. TINJAUAN PUSTAKA

Penelitian mengenai implementasi algoritma C4.5 telah diteliti oleh **Joko Purnomo.dkk** (ISSN : 2338-4018) dengan judul “*Implementasi Algoritma C4.5 dalam Pembuatan Aplikasi Penunjang Keputusan Penerimaan CV. Dinamika Ilmu*”. Sebuah perusahaan penerbitan dan percetakan perlu

mengadakan penerimaan karyawan secara mandiri dilingkungan perusahaan. Kendala yang sering ditemukan dalam proses penerimaan karyawan di sebuah perusahaan yaitu sulitnya menentukan pelamar yang memenuhi kriteria untuk menjadi karyawan sesuai keinginan perusahaan dari beberapa orang pelamar, dan pelamar yang diterima menjadi karyawan terbatas.

Sujana (2010) dengan judul “ *Aplikasi Mining Data Mahasiswa dengan Metode Klasifikasi Decision Tree*”. Data mining merupakan proses analisa data untuk menemukan suatu pola dari kumpulan data tersebut. Data mining mampu menganalisa data yang besar menjadi informasi berupa pola yang mempunyai arti bagi pendukung keputusan. Salah satu teknik yang ada pada data mining adalah klasifikasi. Pada paper ini dibahas teknik klasifikasi yang diterapkan untuk menemukan pola yang terjadi pada data mata kuliah mahasiswa. Teknik klasifikasi yang akan digunakan adalah *Decision tree* , yaitu algoritma C4.5.

Liliana Swastina (2013) dengan judul “ *Penerapan Algoritma C4.5 untuk Penentuan Jurusan Mahasiswa* ”. Banyak kasus dijumpai bahwa pemilihan jurusan yang tidak sesuai dengan kemampuan, kepribadian, minat dan bakat dapat mempengaruhi mahasiswa dalam mengikuti perkuliahan. Penggunaan pendekatan algoritma klasifikasi data mining akan diterapkan untuk menentukan jurusan dalam bidang studi yang akan diambil oleh mahasiswa, sehingga mahasiswa tidak salah dalam memilih jurusan yang akan di tempuh selama belajar pada perguruan tinggi. Algoritma C4.5 digunakan untuk menentukan jurusan yang akan diambil oleh mahasiswa sesuai dengan latar belakang, minat dan kemampuannya sendiri. Parameter pemilihan jurusan adalah Indeks Prestasi Kumulatif Semester 1 dan 2 . Hasil eksperimen dan evaluasi menunjukkan bahwa Algoritma Decision Tree C4.5 akurat diterapkan untuk penentuan kesesuaian jurusan mahasiswa dengan tingkat akurasi 93,31 % dan akurasi rekomendasi jurusan sebesar 82,64%.

Anik Andriani (2012) dengan judul “ Penerapan Algoritma C4.5 pada Program Klasifikasi Mahasiswa Dropout “. Prestasi akademik mahasiswa dievaluasi setiap akhir semester untuk mengetahui hasil belajar yang telah dicapai. Apabila mahasiswa tidak dapat memenuhi kriteria akademik tertentu untuk dinyatakan layak melanjutkan studi maka mahasiswa tersebut dinyatakan putus kuliah atau drop out (DO). Salah satu faktor penyebab banyaknya jumlah mahasiswa DO karena kurangnya kebijakan dan tindakan dari instansi pendidikan untuk menjaga mahasiswa dari DO.

III. METODE PENELITIAN

Adapun perhitungan dalam algoritma c4.5 ini melewati beberapa tahapan sebagai berikut :

Seleksi Data. Sebelum masuk ke proses Data Praproses sing, yang harus dilakukan lebih awal adalah pemilihan data (*data selection*). Data nasabah tersebut nantinya akan menjadi Data Kasus dalam proses operasional Algoritma C 4.5. Dari data yang ada, kolom yang diambil sebagai atribut atau variabel keputusan adalah kolom **Keputusan**, sedangkan kolom yang diambil sebagai variabel penentu dalam pembentukan pohon keputusan adalah kolom : (1) Home Status, (2) Omset Perbulan, (3) Total Pengeluaran, (4) Laba Bersih, (5) OTR, (6) DP, dan (7) Mailing City

Pemilihan variabel-variabel tersebut dengan pertimbangan bahwa jumlah nilai variabelnya tidak banyak sehingga diharapkan calon pegawai yang masuk dalam satu klasifikasi nilai variabel tersebut cukup banyak.

1. Mengelompokkan Home Status

Pengelompokkan nilai *home status* dilakukan dengan memasukkan *home status* calon nasabah dalam *range (jarak)* seperti tampak pada Tabel 4.2 berikut ini.

Home Status	Klasifikasi
Sendiri	Baik Sekali
Keluarga	Baik
Perusahaan	Cukup
Kontrak	Kurang

Tabel 1. Pengelompokkan nilai *Home Status*

2. Mengelompokkan Nominal-Nominal Omset Perbulan

Pengelompokkan Nominal - Nominal (*untuk Omset Perbulan*) dikerjakan dengan memasukkan *Nominal Nominal (untuk Omset Perbulan)* calon nasabah dalam *range (jarak)* seperti tampak pada Tabel 4.3 berikut ini.

Omset Perbulan	Klasifikasi
295.000.001 - 440.000.000	Baik sekali
150.000.001 – 295.000.000	Baik
5.000.0001 – 150.000.000	Cukup
<5.000.000	Kurang

Tabel 4.3 Pengelompokkan Nominal Omset perbulan

3. Mengelompokkan Nominal - Nominal Total Pengeluaran

Pengelompokkan Nominal - Nominal (*untuk Total Pengeluaran*) dikerjakan dengan memasukkan *Nominal Nominal (untuk Total Pengeluaran)* calon Nasabah dalam *range (jarak)* seperti tampak pada Tabel 4.4 berikut ini.

Total Pengeluaran	Klasifikasi
295.000.000 – 440.000.000	Terlalu Tinggi
150.000.001 – 295.000.000	Cukup

5.000.0001 – 150.000.000	Baik
<5.000.000	Baik Sekali

Tabel 4.4 Pengelompokkan Nominal-Nominal Total Pengeluaran

4. Mengelompokkan Nominal - Nominal Laba Bersih

Pengelompokkan Nominal - Nominal (*untuk Laba Bersih*) dikerjakan dengan memasukkan *Nominal Nominal (untuk Laba Bersih)* calon Nasabah dalam *range (jarak)* seperti tampak pada Tabel 4.5 berikut ini.

Lab Bersih	Klasifikasi
295.000.000 – 440.000.000	Baik Sekali
150.000.001 – 295.000.000	Baik
5.000.0001 – 150.000.000	Cukup
<5.000.000	Kurang

Tabel 4.5 Pengelompokkan Nominal - Nominal Laba Bersih

5. Mengelompokkan Nominal - Nominal OTR

Pengelompokkan Nominal - Nominal (*untuk OTR*) dilakukan dengan memasukkan *Nominal - Nominal (untuk OTR)* calon Nasabah dalam *range (jarak)* seperti tampak pada Tabel 4.6 berikut ini.

OTR	Klasifikasi
295.000.000 – 440.000.000	Terlalu Tinggi
150.000.001 – 295.000.000	Tinggi
5.000.0001 – 150.000.000	Cukup
<5.000.000	Tidak Bisa

Tabel 4.6 Pengelompokkan Nominal - Nominal OTR

6. Mengelompokkan Nominal - Nominal DP

Pengelompokkan Nominal - Nominal (*untuk DP*) dilakukan dengan memasukkan *Nominal - Nominal (untuk DP)* calon Nasabah dalam *range (jarak)* seperti tampak pada Tabel 4.7 berikut ini

DP	Klasifikasi
295.000.000 – 440.000.000	Baik Sekali
150.000.001 – 295.000.000	Baik
5.000.0001 – 150.000.000	Cukup
<5.000.000	Kurang

Tabel 4.7 Pengelompokkan Nominal - Nominal DP

IV. HASIL DAN PEMBAHASAN

Proses Perhitungan Algoritma C4.5

Dari proses *Data Preprocessing* yang telah dikerjakan maka dihasilkan tabel yang siap untuk proses klasifikasi Berikut ini adalah penjelasan lebih terperinci mengenai tiap-tiap langkah dalam pembentukan *node* pohon keputusan *tabel kasus* diatas dengan menggunakan algoritma C4.5.

Menghitung jumlah kasus, jumlah kasus untuk keputusan *Lancar*, jumlah kasus untuk keputusan *Tidak Lancar* , dan *Entropy* dari semua kasus dan kasus yang dibagi berdasarkan variabel *Home Status, Omset Perbulan, Total Pengeluaran, Laba Bersih, OTR, DP, Mailing City*.

Secara umum algoritma C4.5 untuk membangun pohon keputusan adalah sebagai berikut:

1. Pilih variabel sebagai akar.
2. Buat cabang untuk tiap-tiap nilai.
3. Bagi kasus dalam cabang.

4. Ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama.

Untuk memilih variabel sebagai akar, didasarkan pada nilai *Gaintertinggi* dari variabel-variabel. Untuk menghitung *Gain* digunakan rumus seperti tertera dalam persamaan (2.1). Setelah mendapatkan nilai *Gain*, ada satu hal lagi yang perlu kita kerjakan perhitungan yaitu mencari nilai *Entropy*. *Entropy* digunakan untuk menentukan seberapa informatif sebuah input variabel untuk menghasilkan output variabel. Untuk menghitung *Entropy* digunakan rumus seperti tertera dalam persamaan (2.2).

Berikut ini adalah penjelasan lebih terperinci mengenai tiap-tiap langkah dalam pembentukan *node* pohon keputusan *tabelkasus* diatas dengan menggunakan algoritma C4.5.

1. Menghitung jumlah kasus, jumlah kasus untuk keputusan **Lancar**, jumlah kasus untuk keputusan **Tidak Lancar**, dan *Entropy* dari semua kasus dan kasus yang dibagi berdasarkan variabel **Home Status, Omset Perbulan, Total Pengeluaran, Laba Bersih, OTR, DP, Mailing City**. Proses perhitungan untuk mendapatkan nilai *Entropy* dan *Gain* setiap variabel adalah sebagai berikut.

a. Menghitung *Entropy* *node* 1

$$Entropy(Total) = \left(-\frac{50}{100} \log_2 \left(\frac{50}{100} \right) + \left(-\frac{50}{100} \log_2 \left(\frac{50}{100} \right) \right) \right) = 1$$

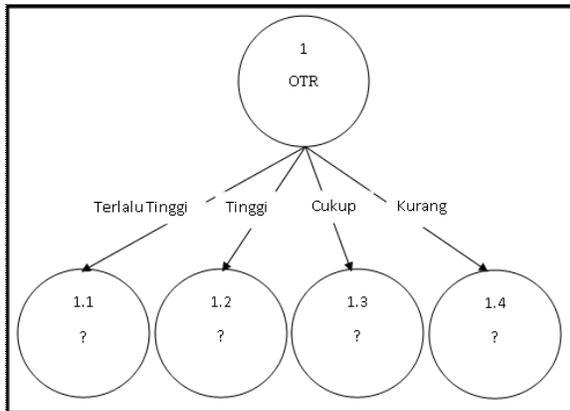
- *Entropy* Variabel Home_Status

$$Entropy(Baik Sekali) = \left(-\frac{46}{92} \log_2 \left(\frac{46}{92} \right) + \left(-\frac{46}{92} \log_2 \left(\frac{46}{92} \right) \right) \right) = 1$$

$$Entropy(Baik) = \left(-\frac{4}{7} \log_2 \left(\frac{4}{7} \right) + \left(-\frac{3}{7} \log_2 \left(\frac{3}{7} \right) \right) \right) = 0.9852$$

$$Entropy(Cukup) = \left(-\frac{0}{1} \log_2 \left(\frac{0}{1} \right) + \left(-\frac{1}{1} \log_2 \left(\frac{1}{1} \right) \right) \right) = 0$$

$$Entropy(Kurang) = \left(-\frac{0}{0} \log_2 \left(\frac{0}{0} \right) + \left(-\frac{0}{0} \log_2 \left(\frac{0}{0} \right) \right) \right) = 0$$



Gambar 4.1 Pohon Keputusan Hasil Perhitungan Node 1

Dengan terbentuknya pohon maka terbentuknya aturan sebagai berikut:

1. Jika otr = Cukup
 - Dan total_pengeluaran = Baik
 - Dan mailing_city = LEBAK
 - Dan home_status = Baik Sekali
 - Dan omset_perbulan = Cukup
 - Dan laba_bersih = Cukup
 - Dan dp = Cukup
 - Maka Tidak Terklasifikasi
2. Jika otr = Cukup
 - Dan total_pengeluaran = Baik
 - Dan mailing_city = PANDEGLANG
 - Dan home_status = Baik Sekali

3. Jika otr = Cukup
 - Dan total_pengeluaran = Baik
 - Dan mailing_city = PANDEGLANG
 - Dan home_status = Baik
 - Maka Lancar
4. Jika otr = Cukup
 - Dan total_pengeluaran = Cukup
 - Maka Lancar
5. Jika otr = Cukup
 - Dan total_pengeluaran = Baik Sekali
 - Dan home_status = Baik Sekali
 - Dan laba_bersih = Cukup
 - Dan mailing_city = LEBAK
 - Dan omset_perbulan = Cukup
 - Dan dp = Cukup
 - Maka Tidak Terklasifikasi
6. Jika otr = Cukup
 - Dan total_pengeluaran = Baik Sekali
 - Dan home_status = Baik Sekali
 - Dan laba_bersih = Cukup
 - Dan mailing_city = PANDEGLANG
 - Dan omset_perbulan = Cukup
 - Dan dp = Cukup
 - Maka Tidak Terklasifikasi
7. Jika otr = Cukup
 - Dan total_pengeluaran = Baik Sekali
 - Dan home_status = Baik Sekali
 - Dan laba_bersih = Kurang
 - Maka Tidak Lancar
8. Jika otr = Cukup
 - Dan total_pengeluaran = Baik Sekali
 - Dan home_status = Baik
 - Dan mailing_city = LEBAK
 - Maka Lancar
9. Jika otr = Cukup
 - Dan total_pengeluaran = Baik sekali
 - Dan home_status = Baik
 - Dan mailing_city = PANDEGLANG
 - Dan omset_perbulan = Cukup
 - Dan laba_bersih = Cukup
 - Dan dp = Cukup
 - Maka Tidak terklasifikasi
10. Jika otr = Cukup
 - Dan total_pengeluaran = Baik Sekali
 - Dan home_status = Cukup
 - Maka Tidak Lancar
11. Jika otr = Tinggi
 - Dan total_pengeluaran = Baik
 - Dan home_status = Baik Sekali
 - Dan omset_perbulan = Cukup
 - Dan laba_bersih = Cukup
 - Dan dp = Cukup
 - Dan mailing_city = LEBAK
 - Maka Tidak Terklasifikasi
12. Jika otr = Tinggi
 - Dan total_pengeluaran = Baik
 - Dan home_status = Baik Sekali
 - Dan omset_perbulan = Cukup

- Dan laba_bersih = Cukup
 Dan dp = Cukup
 Dan mailing_city = PANDEGLANG
 Maka Tidak Terklasifikasi
13. Jika otr = Tinggi
 Dan total_pengeluaran = Baik
 Dan home_status = Baik
 Maka Tidak Lancar
14. Jika otr = Tinggi
 Dan total_pengeluaran = Cukup
 Maka Tidak Lancar
15. Jika otr = Tinggi
 Dan total_pengeluaran = Baik Sekali
 Maka Tidak Lancar

V. SIMPULAN

Setelah perancangan sistem tersebut, maka dapat disimpulkan :

- a. Algoritma C4.5 dapat diimplementasikan dalam pembuatan aplikasi implementasi algoritma c4.5 untuk pengklasifikasian calon nasabah pembiayaan pada PT.Sinar Mitra Sepadan Finace.
- b. Pembentukan pohon keputusan untuk mengklasifikasikan calon nasabah baru dengan variabel tujuan lancar atau tidak lancar dibentuk dengan algoritma C4.5.

VI. DAFTAR PUSTAKA

- [1]. Khantardzic, Mehmed. (2003). *Data Mining : Concepts, Models, Methods and Algorithms, Sons.*
- [2]. Kusriani, Luthfy, E.T.(2009). “*Algoritma Data Mining*”, Yogyakarta : Andi Offset.
- [3]. Purnomo, Joko, et al. “*Implementasi Algoritma C4.5 Dalam Pembuatan Aplikasi Penunjang Keputusan Penerimaan Pegawai CV. Dinamika Ilmu.*” [Online], Jurnal TIKomSiN. ISSN : 2338 – 4018. Tersedia : <http://p3m.sinus.ac.id/jurnal/index.php/TIKomSiN/article/download/158/26/16> September 2014]
- [4]. Sujana. (2010). “Aplikasi Mining Data Mahasiswa Dengan Metode Klasifikasi Decission Tree.”[Online]. ISSN : 1907 - 5022. Tersedia : <http://repository.widyatama.ac.id/xmlui/bitstream/handle/123456789/1277/content.pdf?sequence=116> September 2014].
- [5]. Sujana. (2010). “Klasifikasi Data Nasabah Sebuah Asuransi Menggunakan Algoritma C4.5.” [Online]. ISSN : 1907 - 5022. Tersedia : <http://repository.widyatama.ac.id/xmlui/bitstream/handle/123456789/1276/content.pdf?sequence=116> September 2014].
- [6]. SMSFINANCE, 2014, Profil SMS “Perusahaan Pembiayaan Pilihan Anda”[Online], Tersedia : www.smsfinance.co.id/16 September 2014]