

Implementasi Data Mining Menggunakan Algoritma C4.5 Untuk Menganalisa Resiko Kredit Pada PT Permodalan Nasional Madani

****Desy**, Yohanni Syahra, S.Si., M.Kom. **Suharsil, S.E., M.M. ****

* Program Studi Mahasiswa, STMIK Triguna Dharma

** Program Studi Dosen Pembimbing, STMIK Triguna Dharma

Article Info

Article history:

Keyword:

Data Mining, Metode Algoritma C4.5, Resiko Kredit PT Permodalan Nasional Madani

ABSTRACT

Kemunculan jasa pinjaman memberikan solusi bagi masyarakat jika membutuhkan dana yang mendesak. Namun hal ini juga disertai dengan pola pembayaran si peminjam yang dapat menyebabkan dampak sulitnya terlunasi pinjaman karena faktor-faktor yang tak terduga. Adapun resiko kredit yang merupakan suatu kerugian yang disebabkan kegagalan pihak si peminjam dalam melunasi utang dan suku bunga utang.

Tujuan penelitian ini adalah untuk menganalisis dan merancang data mining yang dapat membantu pengambilan keputusan suatu analisa terhadap masalah yaitu resiko kredit.

Permasalahan yang dihadapi ialah bagaimana cara menganalisis resiko kredit. Oleh Sebab itu, Data Mining hadir untuk memberikan solusi. Hasil dengan penelitian ini dapat membangun sebuah aplikasi sistem data mining untuk menganalisis resiko kredit melalui data yang sudah ada. Dengan adanya data ini diharapkan dapat menyelesaikan permasalahan dalam pengambilan analisis untuk menentukan resiko kredit yang telah terjadi.

Copyright © 2020 STMIK Triguna Dharma.
All rights reserved.

Corresponding Author:

Nama : Desy
Kampus : STMIK Triguna Dharma
Program Studi : Sistem Informasi
E-Mail : deshyass@gmail.com

1. PENDAHULUAN

Kredit merupakan suatu jasa keuangan yang memberikan fasilitas pinjaman uang kepada seseorang atau badan usaha untuk memberikan pinjaman uang dan mewajibkan kepada pihak si peminjam untuk melunasi utang dalam jangka waktu yang telah di tentukan. Pemberian pinjaman kepada pihak si peminjam mengharuskan untuk membayar utang sekaligus membayar bunga jasa pinjaman. Hal itu dilakukan guna mendapatkan untung atas jasa pinjaman uang kepada pihak si peminjam.

Kemunculan jasa pinjaman memberikan solusi bagi masyarakat jika membutuhkan dana yang mendesak. Namun hal ini juga disertai dengan pola pembayaran si peminjam yang dapat menyebabkan dampak sulitnya terlunasi pinjaman karena faktor-faktor yang tak terduga. Dalam kasus tersebut resiko terbesar di tanggung oleh si pemberi pinjaman yang sudah mempercayai untuk memberikan jasa pinjaman uang namun kepada pihak si peminjam mengalami kegagalan dalam melunasi.

Adapun resiko kredit yang merupakan suatu kerugian yang disebabkan kegagalan pihak si peminjam dalam melunasi utang dan suku bunga utang. Resiko kredit menjadi suatu permasalahan yang besar karena setiap rupiah yang tidak dapat terbayar menjadi macet dan berdampak pada timbulnya selisih dalam laporan laba rugi.

Menjadi sebuah permasalahan yang umum dihadapi pihak pemberi pinjam adalah bagaimana menganalisa resiko kredit pada peminjam melalui data yang telah dikumpulkan sebelumnya. Menganalisa resiko kredit yang akan terjadi sangat penting untuk meminimalisir gagalannya pembayaran utang beserta bunganya. Perencanaan yang dilakukan sebelum memberikan pinjaman supaya pihak pemberi pinjaman tidak dengan mudahnya memberikan pinjaman uang. Apabila analisa dilakukan dalam bagian proses perencanaan

menganalisa resiko kredit terhadap pihak si peminjam maka pihak si pemberi pinjaman akan terbantu dalam meminimalkan resiko kredit tersebut.

2. KAJIAN PUSTAKA

2.1 *Knowledge Discovery in Database (KDD)*

Secara sederhana Data Mining adalah bagaimana menelusuri data yang ada untuk membangun sebuah model, kemudian menggunakan model tersebut agar dapat mengenali pola data yang lain yang tidak berada dalam basis data yang tersimpan.

Defenisi lain diantaranya adalah pembelajaran berbasis induksi (*induction-based learning*) adalah proses pembentukan defenisi-defenisi konsep umum yang dilakukan dengan cara mengobservasi contoh-contoh spesifik dari konsep-konsep yang akan dipelajari. *Knowledge Discovery In Database (KDD)* adalah penerapan metode saintifik pada Data Mining. Dalam konteks ini Data Mining merupakan satu langkah dari proses KDD.

2.1.1 Tahapan Proses KDD

Berikut adalah tahapan-tahapan proses KDD yaitu sebagai berikut :

1. *Data Selection*
2. *Pre-Processing dan Cleaning Data*
3. *Transformation*
4. *Data Mining*
5. *Interpretation / Evaluation*

2.1.2 *Data Mining*

Data Mining adalah proses yang menggunakan teknik statistik, perhitungan, kecerdasan buatan dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai basis data besar. Dalam aplikasinya, *data mining* sebenarnya merupakan bagian dari proses *Knowledge Discovery in Database (KDD)*, bukan sebagai teknologi yang utuh dan berdiri sendiri. *Data Mining* merupakan salah satu bagian langkah yang penting dalam proses KDD terutama berkaitan dengan ekstraksi dan penghitungan pola-pola dari data yang ditelaah.

Istilah *Data Mining* memiliki hakikat sebagai disiplin ilmu yang tujuan utamanya adalah untuk menemukan, menggali, atau menambang pengetahuan dari data atau informasi yang kita miliki. *Data Mining* sering juga disebut sebagai *Knowledge Discovery in Database (KDD)*.

2.2 Algoritma C4.5

C4.5 adalah algoritma klasifikasi data dengan tehnik *decision tree* yang terkenal dan disukai karena memiliki kelebihan. Kelebihan ini misalnya dapat mengolah data numerik dan diskret, dapat menangani nilai atribut yang hilang, menghasilkan aturan-aturan yang mudah diinterpretasikan dan tercepat di antara algoritma yang menggunakan memori utama di komputer.

Karakteristik algoritma C4.5 adalah kelompok algoritma *decision tree*. Algoritma ini mempunyai input berupa *training samples* dan *samples*. *Training samples* merupakan data contoh yang digunakan untuk membangun sebuah *tree* yang telah diuji kebenarannya. Sedangkan *samples* merupakan *field-field* data yang digunakan sebagai parameter dalam klasifikasi data.

2.2.1 *Decision Tree*

Decision Tree (pohon keputusan) adalah metode klasifikasi dan prediksi yang mengubah fakta yang sangat besar menjadi pohon keputusan yang mempresentasikan aturan. Pohon keputusan memadukan antara eksplorasi data pemodelan. Pohon keputusan disebut sebagai struktur yang digunakan untuk membagi data yang besar menjadi himpunan-himpunan yang kecil.

2.2.2 Pembuatan *Decision Tree*

Data yang dinyatakan pada pohon keputusan biasanya berbentuk tabel dengan dengan atribut dan record. Atribut menyatakan suatu parameter yang dibuat sebagai kriteria dalam pembentukan pohon keputusan. Atribut yang menyatakan data solusi per item yang memiliki nilai-nilai yang dinamakan dengan intens, dimana proses pada pohon keputusan adalah mengubah bentuk data menjadi model pohon menjadi rule dan menyederhanakan rule.

Bentuk Algoritma C4.5 adalah sebagai berikut :

Input : Sample Training dan Training, Atribut
Output : Pohon Keputusan

- {
- 01 : Membuat simpul akar untuk pohon keputusan yang akan dibuat
- 02 : If semua sempel positif, berhenti dengan suatu pohon dengan satu simpul akar dan beri tanda (+)
- 03 : If semua sempel negatif, berhenti dengan suatu pohon dengan satu simpul akar dan beri tanda (-)
- 04 : If atribut kosong berhenti dengan label sesuai nilai yang terbanyak pada label training

3. Metodologi Penelitian

3.1 Metodologi Penelitian

Metode Penelitian merupakan sebuah proses atau cara ilmiah dalam mendapatkan data yang akan digunakan untuk menyelesaikan masalah.

Adapun tahapan yang dilakukan dalam penelitian guna untuk menyelesaikan permasalahan yang telah dijelaskan pada Bab sebelumnya termasuk pada bagian latar belakang permasalahan, mencakup pada:

1. Teknik Pengumpulan Data (*Data Collection*)

Dalam teknik pengumpulan data dilakukan dua tahapan diantaranya yaitu :

a. Observasi

Kegiatan observasi dalam penelitian ini dilakukan dengan melakukan tinjauan langsung ke PT. Pemodal Nasional Madani. Di perusahaan tersebut dilakukan analisis masalah yang dihadapi dengan cara mengamati proses pengajuan pendaftaran peminjam sebagai syarat awal yang akan diseleksi oleh PT. Permodal Nasional Madani.

b. Wawancara

Setelah itu dilakukan wawancara kepada sub bagian administrasi yang memiliki peran untuk menentukan kelayakan pemilihan peminjam yang layak untuk mendapatkan pinjaman dana dari PT. Permodal Nasional Madani.

c. Studi Literatur (Studi Kepustakaan)

Di dalam studi literatur, penelitian ini banyak menggunakan jurnal-jurnal baik jurnal nasional, jurnal lokal maupun buku sebagai sumber referensi. Dari komposisi yang ada jumlah literature yang digunakan sebanyak 20 dengan rincian: 10 jurnal nasional, 7 jurnal lokal, dan 3 buku nasional. Diharapkan dengan literatur tersebut dapat membantu peneliti di dalam menyelesaikan permasalahan yang terjadi di PT Pemodal Nasional Madani.

3.2 Algoritma Sistem

Algoritma sistem merupakan penjelasan langkah-langkah penyelesaian masalah dalam perancangan sistem penerapan *data mining* dalam menentukan pola kombinasi dengan menggunakan algoritma C4.5. Hal ini dilakukan untuk meningkatkan penilaian kelayakan peminjam sehingga dapat menentukan peminjam yang layak atau tidak.

Algoritma C4.5 dengan rumus :

$$\text{Entropy}(X) = \sum p1 - p1 * \log_2 p1$$

Dimana

X : Himpunan Kasus

m : Jumlah partisi variabel tujuan dari S p1

: Probabilitas kasus dalam partisi ke-i

$$\text{Gain}(S,A) = \text{Entropy}(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * \text{Entropy}(S_i)$$

Dimana

S : Himpunan kasus

A : Variabel penentu

N : Jumlah partisi atribut

A |Si| : Jumlah kasus pada partisi ke-i

|S| : Jumlah kasus dalam S

Langkah awal algoritma C4.5 adalah mencari nilai *entropy*, *entropy* digunakan untuk menentukan seberapa informatif sebuah masukan atribut untuk menghasilkan sebuah atribut. Langkah pertama menentukan nilai *entropy* total dengan jumlah 33 yang terdiri dari Par Meninggal berjumlah 9 dan kelas Par Murni 25. Jumlah kasus setiap atribut akan digunakan untuk menghitung *entropy* total dan *entropy* pada tiap atribut yang ada pada tabel 3.2

Table 3.1 Tabel Perhitungan Nilai *Entropy* dan *Gain* Node 1

Node 1	Atribut	Kategori	Banyak Kasus	Par Meninggal	Par Murni	Entropy	Gain
	Total Kasus		33	9	24	0.8454	
	Jumlah Penghasilan	2-3 Juta	22	4	18	0.6840	
		3-5 Juta	11	5	6	0.9940	
							0.0580
	SaldoUP	<100000	18	2	16	0.5033	
		100000-500000	15	7	8	0.9968	
							0.1178
	Tunggakan Pokok	<100000	3	1	2	0.9183	
		100000-2000000	30	8	22	0.8366	
							0.0013
	Bunga	<100000	19	4	15	0.7425	
		100000-300000	14	5	8	0.9919	
							-0.0029
	Jangka Waktu	<50	28	8	20	0.8631	
		>50	5	1	4	0.7219	
							0.0036

1. Menghitung *Entropy* Total
 $E_{total} (9,25)$
 $=((-9/33)*\text{IMLOG2}(9/33)+(-25/33)*\text{IMLOG2}(25/33))$
 $=0,8454$
2. Perhitungan nilai subset pada atribut penghasilan
 $Entropy (4,18)$
 $=((-4/22)*\text{IMLOG2}(4/22)+(18/22)*\text{IMLOG2}(18/22))$
 $=-0,6840$
 $Entropy (5,6)$
 $=((-4/11)*\text{IMLOG2}(4/11)+(-6/11)*\text{IMLOG2}(6/11))$
 $=0,9940$
3. Perhitungan nilai subset pada atribut saldoup
 $Entropy (2,16)$
 $=((-2/18)*\text{IMLOG2}(2/18)+(-16/18)*\text{IMLOG2}(16/18))$
 $=0,5033$
 $Entropy (5,6)$
 $=((-7/15)*\text{IMLOG2}(7/15)+(-8/15)*\text{IMLOG2}(8/15))$
 $=0,9968$
4. Perhitungan nilai subset pada atribut tunggakan pokok
 $Entropy (1,2)$
 $=((-1/3)*\text{IMLOG2}(1/3)+(-2/3)*\text{IMLOG2}(2/3))$
 $= 0,9183$
 $Entropy (8,22)$
 $=((-8/30)*\text{IMLOG2}(8/30)+(-22/30)*\text{IMLOG2}(22/30))$
 $= 0.8366$
5. Perhitungan nilai subset pada atribut bunga
 $Entropy (4,15)$
 $=((-4/19)*\text{IMLOG2}(4/19)+(-15/19)*\text{IMLOG2}(15/19))$
 $= 0,7425$
 $Entropy (5,8)$

$$= ((-5/14)*\text{IMLOG2}(5/14)+(-8/14)*\text{IMLOG2}(8/14))$$

$$= 0,9919$$

6. Perhitungan nilai subset pada atribut jangka waktu

$$\text{Entropy}(8,20)$$

$$= ((-8/28)*\text{IMLOG2}(8/28)+(-20/28)*\text{IMLOG2}(20/28))$$

$$= 0,8631$$

$$\text{Entropy}(1,4)$$

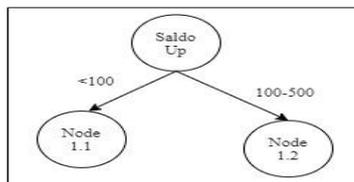
$$= ((-1/5)*\text{IMLOG2}(1/5)+(-4/5)*\text{IMLOG2}(4/5))$$

$$= 0,7219$$

Setelah melakukan semua perhitungan nilai *entropy* pada masing-masing atribut, maka tahap berikutnya tahap berikutnya adalah menghitung nilai *gain*.

1. Perhitungan Nilai *Gain* Pada Atribut Penghasilan
 $= (0,8454)-((22/33)*0,6840)-((11/33)*0,9940)$
 $= 0,0580$
2. Perhitungan Nilai *Gain* Pada Atribut SaldoUp
 $= (0,8454)-((18/33)*0,5033)-((15/33)*0,9968)$
 $= 0,1178$
3. Perhitungan Nilai *Gain* Pada Atribut Tunggalan Pokok
 $= (0,8454)-((3/33)*0,9183)-((30/33)*0,8366)$
 $= 0,0013$
4. Perhitungan Nilai *Gain* Pada Atribut Bunga
 $= (0,8454)-((19/33)*0,7425)-((14/33)*0,9919)$
 $= -0,0029$
5. Perhitungan Nilai *Gain* Pada Atribut Jangka Waktu
 $= (0,8454)-((28/33)*0,8631)-((5/33)*0,7219)$
 $= 0,0036$

Karena *gain* terbesar adalah *saldoup*, *saldoup* menjadi node akar(*root node*). Setelah selesai menghitung nilai *gain* tiap atribut, maka selanjutnya adalah membuat pohon keputusan. Seperti dibawah ini :



Gambar 3.1 Pohon Keputusan Hasil Perhitungan *Node 1*

Untuk mengetahui cabang berikutnya dilakukan perhitungan jumlah kasus untuk keputusan Par Meninggal dan jumlah keputusan Par Murni dan *entropy* nya dari semua kasus. Kasus yang dibagi berdasarkan atribut penghasilan, tunggakan pokok, bunga dan jangka waktu yang dapat menjadi akar dari nilai. Setelah itu dilakukan perhitungan *gain* untuk tiap- tiap atribut, hasil perhitungan ada dibawah ini :

Tabel 3.2 Tabel Perhitungan Nilai *Entropy* dan *Gain* *Node1* 1.1

Node 1.1	Atribut	Kategori	Banyak Kasus	Par Meninggal	Par Murni	Entropy	Gain
	Total Kasus		15	7	8	0.9968	
	Jumlah Penghasilan	2-3 Juta	9	4	5	0.9911	
		3-5 Juta	6	3	3	1.0000	
							0.0021
	Tunggakan Pokok	<100000	3	1	2	0.9183	
		100000-200000	12	6	6	1.0000	

							0.0131
	Bunga	<100000	6	4	0	0.0000	
		100000-300000	9	3	6	0.9183	
							0.4458
	Jangka Waktu	<50	14	6	8	0.9852	
		>50	1	1	0	0.0000	
							0.0772

- Perhitungan nilai *Entropy* Total

$$=((-7/15)*\text{IMLOG2}(7/15)+(-8/15)*\text{IMLOG2}(8/15))$$

$$=0,9968$$
- Perhitungan nilai subset pada atribut jumlah penghasilan
Entropy (4,5)

$$=((-4/9)*\text{IMLOG2}(4/9)+(-5/9)*\text{IMLOG2}(5/9))$$

$$= 0,9911$$
Entropy (3,3)

$$=((-3/6)*\text{IMLOG2}(3/6)+(-3/6)*\text{IMLOG2}(3/6))$$

$$= 1,0000$$
- Perhitungan nilai subset pada atribut tunggakan pokok
Entropy (1,2)

$$=((-1/3)*\text{IMLOG2}(1/3)+(-2/3)*\text{IMLOG2}(2/3))$$

$$= 0,9183$$
Entropy (6,6)

$$=((-6/12)*\text{IMLOG2}(6/12)+(-6/12)*\text{IMLOG2}(6/12))$$

$$= 1,0000$$
- Perhitungan nilai subset pada atribut bunga
Entropy (4,0)

$$=((-4/6)*\text{IMLOG2}(4/6)+(-0/6)*\text{IMLOG2}(0/6))$$

$$= 0,0000$$
Entropy (3,6)

$$=((-3/9)*\text{IMLOG2}(3/9)+(-6/9)*\text{IMLOG2}(6/9))$$

$$= 0,9183$$
- Perhitungan nilai subset pada atribut jangka waktu
Entropy (6,8)

$$=((-6/14)*\text{IMLOG2}(6/14)+(-8/14)*\text{IMLOG2}(8/14))$$

$$= 0,9852$$
Entropy (1,0)

$$=((-1/1)*\text{IMLOG2}(1/1)+(-0/1)*\text{IMLOG2}(0/1))$$

$$= 0,0000$$

Setelah melakukan semua perhitungan nilai *entropy* pada masing-masing atribut, maka tahap berikutnya adalah melakukan perhitungan nilai *gain*.

- Perhitungan nilai *gain* pada atribut jumlah penghasilan

$$= (0,9968)-((9/15)* 0,9911)-((6/15)*1,0000)$$

$$= 0.0021$$
- Perhitungan nilai *gain* pada atribut tunggakan pokok

$$=(0,9968)-((3/15)* 0,9183)-((12/15)*1,0000)$$

$$=1$$
- Perhitungan nilai *gain* pada atribut bunga

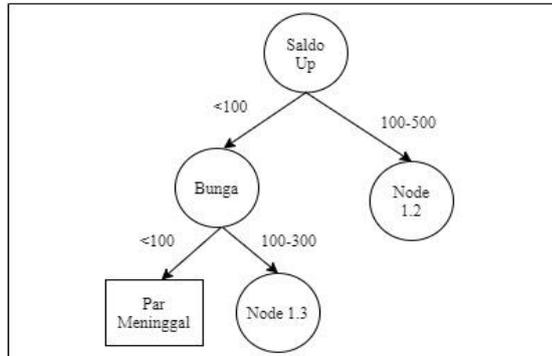
$$=(0,9968)-((6/15)*0.0000)-((9/15)*0,9183)$$

$$=0,4458$$
- Perhitungan nilai *gain* pada atribut bunga

$$=(0,9968)-((14/15)*0,9852)-((1/15)*0,0000)$$

=0,0772

Karena *gain* terbesar adalah bunga, bunga menjadi node akar(*root node*). Setelah selesai menghitung nilai *gain* tiap atribut, maka selanjutnya adalah membuat pohon keputusan. Seperti dibawah ini :



Gambar 3.2 Pohon Keputusan Hasil Perhitungan *Node 1.1*

Untuk mengetahui cabang berikutnya dilakukan perhitungan jumlah kasus untuk keputusan Par Meninggal dan jumlah keputusan Par Murni dan entropy nya dari semua kasus. Kasus yang dibagi berdasarkan atribut penghasilan, tunggakan pokok, jangka waktu yang dapat menjadi akar dari nilai. Setelah itu dilakukan perhitungan *gain* untuk tiap- tiap atribut, hasil perhitungan ada dibawah ini :

Tabel 3.3 Tabel Perhitungan Nilai *Entropy* dan *Gain Node 1.2*

Node 1.2	Atribut	Kategori	Banyak Kasus	Par Meninggal	Par Murni	Entropy	Gain
	Total Kasus		9	3	6	0.9183	
	Jumlah Penghasilan	2-3 Juta	4	3	1	0.8113	
		3-5 Juta	5	0	5	0.0000	
							0.5577
	Tunggakan Pokok	<100000	0	0	0	0.0000	
		100000-2000000	9	3	6	0.9183	
							0.0000
	Jangka Waktu	<50	8	2	6	0.8113	
		>50	1	1	0	0.0000	
							0.1972

- Perhitungan nilai *Entropy* Total
 $=((-3/9)*\text{IMLOG2}(3/9)+(-6/9)*\text{IMLOG2}(6/9))$
 $=0.9183$
- Perhitungan nilai subset pada atribut jumlah penghasilan
Entropy (3,1)
 $=((-3/4)*\text{IMLOG2}(3/4)+(-1/4)*\text{IMLOG2}(1/4))$
 $=0,8113$
Entropy (0,5)
 $=((-0/5)*\text{IMLOG2}(0/5)+(-5/5)*\text{IMLOG2}(5/5))$
 $=0,0000$
- Perhitungan nilai subset pada atribut tunggakan pokok
Entropy (0,0)
 $=((-0/0)*\text{IMLOG2}(0/0)+(-0/0)*\text{IMLOG2}(0/0))$
 $= 0.0000$
Entropy (3,6)

$$= ((-3/9)*\text{IMLOG2}(3/9)+(-6/9)*\text{IMLOG2}(6/9))$$

$$= 0,9183$$

4. Perhitungan nilai subset pada atribut jangka waktu

Entropy (2,6)

$$= ((-2/8)*\text{IMLOG2}(2/8)+(-6/8)*\text{IMLOG2}(6/8))$$

$$= 0.8113$$

Entropy (1,0)

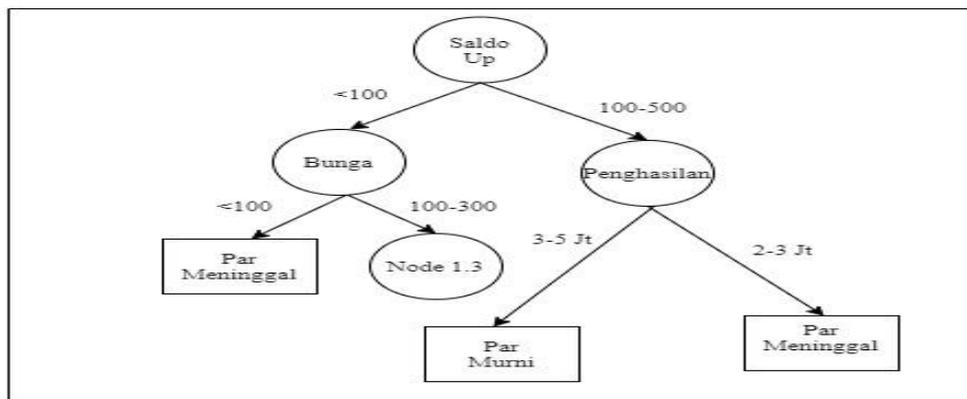
$$= ((-1/1)*\text{IMLOG2}(1/1)+(-0/0)*\text{IMLOG2}(0/0))$$

$$= 0,0000$$

Setelah melakukan semua perhitungan nilai *entropy* pada masing-masing atribut, maka tahap berikutnya adalah melakukan perhitungan nilai *gain*.

1. Perhitungan nilai *gain* pada atribut penghasilan
 $= (0,9183) - ((4/9) * 0,8113) - ((5/9) * 0,0000)$
 $= 0,5577$
2. Perhitungan nilai *gain* pada atribut tunggakan pokok
 $= (0,9183) - ((0/9) * 0,0000) - ((9/9) * 0,9183)$
 $= 0,0000$
3. Perhitungan nilai *gain* pada atribut jangka waktu
 $= (0,9183) - ((8/9) * 0,8113) - ((1/9) * 0,0000)$
 $= 0,1972$

Karena *gain* terbesar adalah penghasilan, penghasilan menjadi node akar (*root node*). Setelah selesai menghitung nilai *gain* tiap atribut, maka selanjutnya adalah membuat pohon keputusan. Seperti dibawah ini



Gambar 3.3 Pohon Keputusan Hasil Perhitungan *Node 1.2*

Untuk mengetahui cabang berikutnya dilakukan perhitungan jumlah kasus untuk keputusan Par Meninggal dan jumlah keputusan Par Murni dan *entropy* nya dari semua kasus. Kasus yang dibagi berdasarkan atribut Tunggakan Pokok, Jangka Waktu yang dapat menjadi akar dari nilai. Setelah itu dilakukan perhitungan *gain* untuk tiap- tiap atribut, hasil perhitungan ada dibawah ini :

Tabel 3.4 Tabel Perhitungan Nilai *Entropy* dan *Gain Node 1.3*

Node 1.3	Atribut	Kategori	Banyak Kasus	Par Meninggal	Par Murni	Entropy	Gain
	Total Kasus		4	3	1	0.8113	
	Tunggakan Pokok	<100000	0	0	0	0.0000	
		100000-2000000	4	3	1	0.8113	0.0000

	Jangka Waktu	<50	4	1	3	0.8113	
		>50	0	0	0	0.0000	
							0.0000

1. Perhitungan nilai *Entropy* Total

$$=((-3/4)*\text{IMLOG2}(3/4)+(-1/4)*\text{IMLOG2}(1/4))$$

$$=0,8113$$
2. Perhitungan nilai subset pada atribut tunggakan pokok
Entropy (0,0)

$$=((-0/0)*\text{IMLOG2}(0/0)+(-0/0)*\text{IMLOG2}(0/0))$$

$$=0,000$$
Entropy (3,1)

$$=((-3/4)*\text{IMLOG2}(3/4)+(-1/4)*\text{IMLOG2}(1/4))$$

$$=0,8113$$
3. Perhitungan nilai subset pada atribut jangka waktu
Entropy (1,3)

$$=((-1/4)*\text{IMLOG2}(1/4)+(-3/4)*\text{IMLOG2}(3/4))$$

$$=0,8113$$
Entropy (0,0)

$$=((-0/0)*\text{IMLOG2}(0/0)+(-0/0)*\text{IMLOG2}(0/0))$$

$$=0,0000$$

Setelah melakukan semua perhitungan nilai *entropy* pada masing-masing atribut, maka tahap berikutnya adalah melakukan perhitungan nilai *gain*.

1. Perhitungan nilai *gain* pada atribut tunggakan pokok

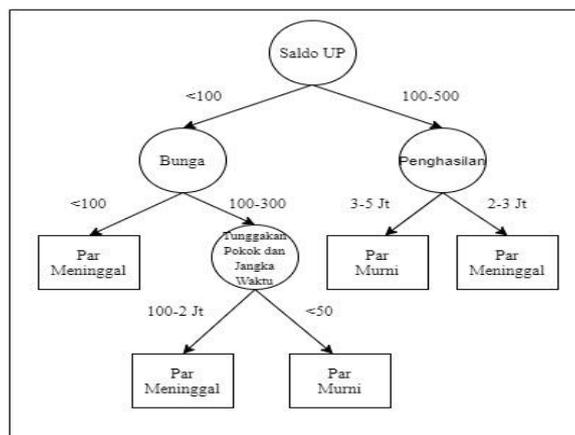
$$=(0,8113)-((0/4)* 0,000)-((4/4)*0,8113)$$

$$=0,000$$
2. Perhitungan nilai *gain* pada atribut jangka waktu

$$=(0,8113)-((4/4)* 0,8113)-((0/4)*0,000)$$

$$=0,0000$$

Karena *gain* terbesar adalah Sama, maka kedua nilai menjadi node akar(*root node*). Setelah selesai menghitung nilai *gain* tiap atribut, maka selanjutnya adalah membuat pohon keputusan. Seperti dibawah ini :



Gambar 3.4 Pohon Keputusan Hasil Perhitungan *Node* 1.3

Dengan memperhatikan pohon keputusan pada gambar di atas diketahui bahwa pohon keputusan telah terbentuk. Dan setelah didapatkan *tree*, maka diperoleh *rule*. Berikut adalah bentuk *tree* yang diubah menjadi *rule* :

1. *If* Saldo Up = "<100" and Bunga = "<100" then Status ="Par Meninggal"
2. *If* Saldo Up = "<100" and Bunga = "100-300" and Tunggakan Pokok dan Jangka Waktu ="100-2Jt" then Status ="Par Mninggal" and ="<50" then Status ="Par Murni"
3. *If* Saldo Up = "<100" and Bunga = "<100" then Status ="Par Murni"

4. *If Saldo Up = "100-500" and Penghasilan = "2-3 Jt" then Status = "Par Meninggal" and Penghasilan = "3-5 Jt" then Status = "Par Murni"*

4 PENGUJIAN DAN IMPLEMENTASI

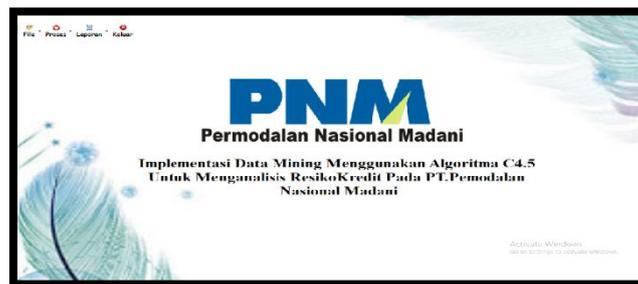
1. Form Login

Berikut ini merupakan tampilan dari *Form* login yang berfungsi untuk melakukan proses validasi username dan password pengguna sebelum masuk kedalam Menu Utama:

Gambar 5.1 *Form* Login

2. Form Menu Utama

Menu utama merupakan tampilan yang akan muncul setelah admin melakukan *login* pada aplikasi sistem data *mining* untuk implementasi data mining menggunakan algoritma C4.5 untuk menganalisa resiko kredit pada PT Permodalan Nasional Madani dengan menggunakan algoritma klasifikasi C4.5.



Gambar 5.2 *Form* Menu Utama

3. Form Data Atribut

Form data atribut merupakan tampilan *form* untuk masuk ke data atribut yang digunakan admin untuk dapat melihat data atribut. Tampilan *form* data atribut dapat dilihat pada gambar dibawah ini :

Kode Atribut	Nama Atribut	Nilai Atribut
001	Jumlah Pengha...	2-3 Juta
002	Jumlah Pengha...	3-5 Juta
003	SaldoUP	<100000
004	SaldoUP	100000-500000
005	Tunggakan Po...	<100000
006	Tunggakan Po...	100000-2000000
007	Bunga	<100000

Gambar 5.3 Tampilan *Form* Data Atribut.

4. Form Data Kasus

Form data kasus merupakan tampilan *form* untuk masuk ke data kasus yang digunakan admin untuk dapat melihat data kasus. Tampilan *form* data kasus dapat dilihat pada gambar dibawah ini :

NasabahId	Nama	Tunggakan Pokok	Tunggakan Bunga	SaldoUP	Penghasilan	Jangka Waktu	Status Nasabah
90365000274	Iin Binti A...	100000-2000000	100000-3000000	<100000	2-3 Juta	<=50	Par Murni
91785002259	BONTOR ...	100000-2000000	<100000	<100000	2-3 Juta	<=50	Par Murni
90365000431	DELIANA ...	100000-2000000	100000-3000000	100000-50000...	3-5 Juta	<=50	Par Murni
90366002190	Leginem ...	100000-2000000	100000-3000000	100000-50000...	3-5 Juta	<=50	Par Meningga
90366000860	LINDA WA...	100000-2000000	<100000	100000-50000...	2-3 Juta	<=50	Par Murni
90365006970	Chairani ...	100000-2000000	100000-3000000	100000-50000...	2-3 Juta	<=50	Par Murni
90365005147	RINI SART...	100000-2000000	<100000	<100000	3-5 Juta	<=50	Par Murni
90366004567	DEWILES...	<100000	<100000	100000-50000...	2-3 Juta	<=50	Par Murni

Gambar 5.4 Tampilan Form Data Kasus.

5. Form Data Proses

Form data proses merupakan tampilan form untuk masuk ke data proses yang digunakan admin untuk dapat melihat data proses. Tampilan form data proses dapat dilihat pada gambar dibawah ini :

Nasabah ID	Nama	Tunggakan Pokok	Tunggakan Bunga	Saldo UP	Penghasilan	Jangka Waktu	Status Nas.
90365000...	Iin Binti Ajit	100000-2000000	100000-3000000	<100000	2-3 Juta	<=50	Par Murni
91785002...	BONTOR TATY...	100000-2000000	<100000	<100000	2-3 Juta	<=50	Par Murni
90365000...	DELIANA SIHO...	100000-2000000	100000-3000000	100000-50000...	3-5 Juta	<=50	Par Murni
90366002...	Leginem Binti ...	100000-2000000	100000-3000000	100000-50000...	3-5 Juta	<=50	Par Menin.
90366000...	LINDA WAI BI...	100000-2000000	<100000	100000-50000...	2-3 Juta	<=50	Par Murni
90365006...	Chairani Binti ...	100000-2000000	100000-3000000	100000-50000...	2-3 Juta	<=50	Par Murni

Total Nasabah 33 Par Meninggal 9 Par Murni 24 Entropy Total 0.845350936622437

Gambar 5.5 Tampilan Form Data Proses

6. Form Laporan

Form laporan merupakan tampilan form untuk menampilkan hasil dari pengujian sistem yang digunakan admin untuk dapat melihat hasil laporan. Tampilan form laporan dapat dilihat pada gambar dibawah ini :

PNM
PT. Pemodal Nasional Madani

Laporan Hasil Teknisi Mesin EDC (Electronic Data Capture) yang layak diterima
Pada PT. SOUCI INDOPRIMA Dengan Menggunakan Metode Oreste

NasabahId	Nama	Tunggakan Pokok	Tunggakan Bunga	SaldoUp	Penghasilan	Jangka Waktu	Status
90365000274	Iin Binti Ajit	100000-3000000	100000-2000000	<100000	2-3 Juta	<50	Par Murni

Gambar 5.6 Tampilan Form Laporan

5. KESIMPULAN

Berdasarkan hasil pembahasan dan pengamatan yang telah dilakukan dari permasalahan yang terjadi tentang kasus Implementasi Data Mining Menggunakan Algoritma C4.5 Untuk Menganalisa Resiko Kredit Pada PT Permodalan Nasional Madani maka dapat diambil kesimpulan diantaranya :

1. Hasil analisa dalam menentukan resiko kredit dapat di deskripsikan dalam bentuk rule ataupun pohon keputusan.

2. Penerapan Algoritma Klasifikasi C4.5 dalam menentukan resiko kredit, dalam sistem ini mampu memberikan hasil resiko kredit berupa nilai, pohon keputusan dan presentase .
3. Perancangan dan pembangun aplikasi data mining menggunakan metode algoritma klasifikasi C4.5, hasil uji sistem dapat dilihat sejauh mana sistem tersebut mampu menangani dalam pemecahan masalah dalam menentukan resiko kredit.

UCAPAN TERIMA KASIH

Puji syukur kehadiran Allah SWT atas izin-Nya yang telah melimpahkan rahmat dan karunia-Nya sehingga dapat menyelesaikan jurnal ilmiah ini. Pada kesempatan ini diucapkan terima kasih yang sebesar-besarnya kepada kedua Orang Tua tercinta yang selama ini memberikan do'a dan dorongan baik secara moril maupun materi sehingga dapat terselesaikan pendidikan dari tingkat dasar sampai bangku perkuliahan dan terselesaikannya jurnal ini. Di dalam penyusunan jurnal ini, banyak sekali bimbingan yang didapatkan serta arahan dan bantuan dari pihak yang sangat mendukung. Oleh karena itu dengan segala kerendahan hati, diucapkan terima kasih yang sebesar-besarnya kepada Bapak Rudi Gunawan, SE., M.Si., selaku Ketua Sekolah Tinggi Manajemen InFormatika Dan Komputer (STMIK) Triguna Dharma Medan. Bapak Dr. Zulfian Azmi, ST., M.Kom., selaku Wakil Ketua I Bidang Akademik STMIK Triguna Dharma Medan. Bapak Marsono, S.Kom., M.Kom., selaku Ketua Program Studi Sistem Informasi STMIK Triguna Dharma Medan. Ibu Yohhani Syahra, S.Si., M.Kom., selaku Dosen Pembimbing I yang telah memberikan saran, arahan dan dukungannya serta motivasi, sehingga penelitian ini dapat terselesaikan dengan baik dan tepat waktu. Bapak Suharsil, ., S.E, M.M., selaku Dosen Pembimbing II yang telah memberikan bimbingan tata cara penulisan, saran dan motivasi sehingga penelitian ini dapat terselesaikan dengan baik dan tepat waktu. Seluruh Dosen, Staff dan Pegawai di STMIK Triguna Dharma Medan.

REFERENSI

- [1] Santoso, T. B., & Sekardiana, D. (2019). Penerapan algoritma c4. 5 untuk penentuan kelayakan pemberian kredit (studi kasus: Koperia-koperasi warga komplek gandaria). *Jurnal Algoritma, Logika dan Komputasi*, 2(1).
- [2] Khasanah, S. N. (2017). Penerapan Algoritma C4. 5 Untuk Penentuan Kelayakan Kredit. *Jurnal Techno Nusa Mandiri*, 14(1), 9-14.
- [3] Mardi, Y. (2017). Data Mining: Klasifikasi Menggunakan Algoritma C4. 5. *Edik Informatika*, 2(2), 213-219.
- [4] Syakur Muhammad Luthfan. (2014). No Title. *Sistem Informasi Penyewaan Lapangan Futsal Pada Grindulu Futsal Pacitan*, 3 No 4, 29–36.

BIOGRAFI PENULIS

	<p>Desy, Perempuan kelahiran Lubuk Pakam, 21 Desember 1996, merupakan seorang mahasiswa STMIK Triguna Dharma yang sedang dalam proses menyelesaikan skripsi.</p>
	<p>Yohanni Syahra, S.Si, M.Kom, Beliau merupakan dosen STMIK Triguna Dharma. Dosen Pembimbing I</p>
	<p>Suharsil, S.E, M.M, Beliau merupakan dosen STMIK Triguna Dharma. Dosen Pembimbing II</p>