

# PENERAPAN ALGORITMA C 4.5 DALAM MEMPEROLEH DECISION TREE UNTUK MEMPREDIKSI PENENTUAN RESIKO KREDIT PADA BANK BPR BUKITTANDANG MANDIRI PADANG MENGUNAKAN ESTARD DATA MINER

**Eka Praja Wiyata Mandala, S.Kom, M.Kom**  
**Teknik Informatika, Universitas Putra Indonesia “YPTK” Padang**

e-mail : eka\_praja\_199@yahoo.com

**Abstrak** – Bank Perkreditan Rakyat Bukit Tandang Mandiri cabang Padang merupakan salah satu Bank yang telah memiliki sistem komputerisasi yang cukup baik, namun untuk penganalisisan resiko kredit data calon debitur masih menggunakan cara manual. Salah satu teknik yang dapat membantu pihak bank dalam penganalisisan resiko kredit adalah teknik decision tree dengan menggunakan algoritma C4.5, yang merupakan bagian dari teknik data mining untuk mengambil keputusan dalam memberikan persetujuan kredit kepada calon debitur yang telah dianalisa. Cara pengolahan data dan penganalisisan tersebut dilakukan dengan menggunakan software Estard Data Miner. Software ini dapat menghasilkan pohon keputusan untuk mengetahui resiko besar kecilnya resiko kredit terhadap calon debitur yang akan diberikan pinjaman

**Kata Kunci:** *Data Mining, Decision Tree, Algoritma C4.5*

## PENDAHULUAN

Saat ini sistem komputerisasi tidak hanya untuk perhitungan berulang, kecepatan waktu serta keakuratan untuk membuat laporan-laporan yang diperlukan sebagai penunjang aktifitas perusahaan, namun juga sampai pada perusahaan yg bergerak dibidang jasa. Salah satu fungsi dari informasi hasil pengolahan datanya adalah dapat menghasilkan informasi yang terkait dengan resiko kredit pada Bank. Untuk mendapatkan informasi tersebut dapat digunakan salah satu metode decision tree dan algoritma C4.5 yang ada pada data mining.

Berdasarkan uraian di atas, dapat dirumuskan beberapa masalah sebagai berikut :

1. Bagaimana sistem yang dianalisis dapat mempermudah pihak Bank BPR dalam mengambil keputusan resiko kredit ?
2. Bagaimana aplikasi data mining yang dibuat, dapat membantu di dalam proses penentuan resiko kredit?

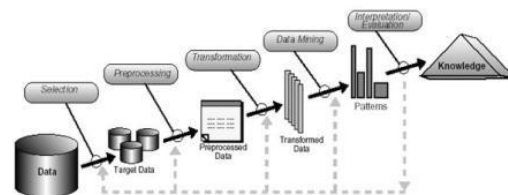
Agar penulisan lebih terarah, permasalahan yang dihadapi maka ditetapkan batasan-batasan terhadap sistem yang diteliti yaitu menerapkan teknik *decision tree* dan algoritma C4.5 untuk menentukan resiko kredit pada Bank BPR Bukittandang Mandiri Padang.

Tujuan pembuatan aplikasi data mining untuk prediksi penentuan resiko kredit pada Bank BPR adalah untuk mempermudah pihak Bank dalam melakukan penentuan resiko kredit, sehingga pihak bank bisa menentukan resiko yang akan terjadi sebelum nasabah melakukan peminjaman.

## LANDASAN TEORI

*Data mining* disebut juga sebagai serangkaian proses untuk menggali nilai tambah berupa pengetahuan yang selama ini tidak diketahui secara manual dari suatu kumpulan data [1].

*Data mining* sering juga disebut sebagai *knowledge discovery in database* (KDD). KDD adalah kegiatan yang meliputi pengumpulan, pemakaian data, historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar [2].



Gambar 1. Proses KDD

Dalam klasifikasi, keluaran dari dari setiap data adalah bilangan bulat atau diskrit. Dalam contoh pengambilan keputusan untuk main bola atau tidak maka keuaran tersebut bisa diubah ke bilangan bulat 1 atau -1. Dengan melihat keluaran yang berupa bilangan bulat kita bisa menerapkan metode klasifikasi [2]

Dalam *decision tree* tidak menggunakan *vector* jarak untuk mengklasifikasikan obyek. Seringkali data observasi mempunyai atribut-atribut yang bernilai nominal. Misalkan obyeknya adalah sekumpulan buah-buahan yang bisa dibedakan berdasarkan atribut bentuk, warna, ukuran dan rasa. Bentuk, warna, ukuran

dan rasa adalah besaran nominal, yaitu bersifat kategoris dan tiap nilai tidak bisa dijumlahkan atau dikurangkan. Dalam atribut warna ada beberapa nilai yang mungkin yaitu hijau, kuning, merah. Dalam atribut ukuran ada nilai besar, sedang dan kecil. Dengan nilai-nilai atribut ini, kemudian dibuat *decision tree* untuk menentukan suatu obyek termasuk jenis buah apa jika nilai tiap-tiap atribut diberikan [2]

*Decision tree* sesuai digunakan untuk kasus-kasus yang keluarannya bernilai diskrit. Walaupun banyak variasi model *decision tree* dengan tingkat kemampuan dan syarat yang berbeda, pada umumnya beberapa ciri yang cocok untuk diterapkannya *decision tree* adalah sebagai berikut [2]

Algoritma c4.5 merupakan algoritma yang digunakan untuk membentuk pohon keputusan [1]

Secara umum algoritma C4.5 untuk membangun pohon keputusan adalah sebagai berikut [1]:

1. Pilih atribut sebagai akar
2. Buat cabang untuk masing-masing nilai
3. Bagi kasus dalam cabang
4. Ulangi proses untuk masing-masing cabang sampai semua kasus pada cabang memiliki kelas yang sama.

Untuk perhitungan nilai *entropy* dapat dilihat pada persamaan sebagai berikut [1]:

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

Keterangan :

- S : himpunan kasus
- A : fitur
- n : jumlah partisi S
- $p_i$  : proporsi dari  $S_i$  terhadap S

Untuk memilih atribut sebagai akar, didasarkan pada nilai gain tertinggi dari atribut-atribut yang ada. Untuk menghitung gain digunakan rumus seperti tertera dalam persamaan berikut [1]:

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

Keterangan :

- S : himpunan kasus
- A : atribut
- n : jumlah partisi atribut A
- $|S_i|$  : jumlah kasus pada partisi ke-i
- $|S|$  : jumlah kasus dalam S

ESTARD Data Miner adalah komprehensif perangkat lunak data mining yang mampu menemukan pengetahuan yang paling tak terduga dalam database. Dibandingkan dengan alat intelijen bisnis umum lainnya, ESTARD Data Miner dapat menyediakan sesuatu yang lebih dari sekedar operasi statistik serta memberi kekuatan untuk bekerja dengan analisis prediktif [4].

ESTARD Data miner digunakan untuk [4]:

- a. Melakukan analisis statistik untuk data dengan pemahaman yang lebih baik.
- b. Mengungkapkan/menemukan korelasi yang tersembunyi dan dependensi di database dengan bantuan aturan keputusan (rule) dan pohon keputusan (decision tree) .
- c. Untuk membuat pola data deskripsi.
- d. Membuat dan menerapkan model keputusan.
- e. Bekerja dengan analisis prediktif

Risiko kredit (*credit risk*) adalah merupakan suatu risiko kerugian yang disebabkan oleh ketidak mampuan (gagal bayar) dari debitur atas kewajiban pembayaran utangnya baik utang pokok maupun bunganya ataupun keduanya [3].

Risiko ini merupakan kondisi dan situasi yang akan dihadapi di masa yang akan datang yang sangat besar pengaruhnya terhadap perolehan laba bank [3]

Hal-hal yang termasuk dalam Risiko Kredit adalah [3]:

1. *Lending Risk*, yaitu risiko akibat nasabah/debitur tidak mampu melunasi fasilitas yang telah diberikan oleh bank, baik berupa fasilitas kredit langsung maupun tidak langsung (*cash loan* maupun *non cash loan*)
2. *Counterparty Risk*, risiko dimana *counterpart* tidak bisa melunasi kewajibannya ke bank baik sebelum tanggal kesepakatan maupun pada saat tanggal kesepakatan.
3. *Issuer Risk*, risiko dimana penerbit suatu surat berharga tidak bisa melunasi kepada bank sejumlah nilai surat berharga yang dimiliki bank.

## HASIL DAN PEMBAHASAN

Dalam mengumpulkan data, digunakan teknik wawancara dan dokumentasi yaitu pertemuan antara pengumpul data dengan Kepala cabang serta karyawan bank untuk bertukar informasi agar data yang diperoleh dapat diolah. Dan sumber datanya merupakan sumber data primer, yaitu data yang diberikan secara langsung.

Tabel 1. Sampel Data dari Bank BPR

No	Tujuan pinjam	Jumlah Pinjam	Jangka Waktu	Bunga	Resiko
1	Modal Kerja	Rp.20.000.000,-	24 Bulan	24%Tahun	Kecil
2	Konsumtif	Rp.50.000.000,-	60 Bulan	25%Tahun	Sedang
3	Investasi	Rp.30.000.000,-	36 Bulan	20%Tahun	Sedang
4	Konsumtif	Rp.3.000.000,-	24 Bulan	24%Tahun	Kecil
5	Konsumtif	Rp.10.000.000,-	12 Bulan	20%Tahun	Kecil
6	Modal Kerja	Rp.2.000.000,-	12 Bulan	24%Tahun	Kecil
7	Konsumtif	Rp.8.000.000,-	24 Bulan	25%Tahun	Kecil
8	Modal Kerja	Rp.2.000.000,-	12 Bulan	24%Tahun	Kecil
9	Investasi	Rp.30.000.000,-	36 Bulan	22%Tahun	Sedang
10	Konsumtif	Rp.3.000.000,-	24 Bulan	24%Tahun	Kecil
11	Investasi	Rp.40.000.000,-	48 Bulan	23%Tahun	Besar
12	Konsumtif	Rp.40.000.000,-	48 Bulan	25%Tahun	Besar
13	Modal Kerja	Rp.20.000.000,-	24 Bulan	23%Tahun	Kecil
14	Modal Kerja	Rp.40.000.000,-	48 Bulan	21%Tahun	Besar
15	Konsumtif	Rp.30.000.000,-	36 Bulan	25%Tahun	Sedang

Dari tabel diatas, dapat dianalisa resiko pemberian kredit yang diberikan kepada 15 orang nasabah Bank BPR. Sehingga diperoleh data hasil analisa seperti dibawah ini.

Tabel 2. Hasil Analisa Penentuan Resiko Kredit

No	Jumlah Pinjam	Jangka Waktu	Suku Bunga	Resiko
1	Pendek	Kecil	Besar	Kecil
2	Panjang	Menengah	Besar	Sedang
3	Menengah	Menengah	Kecil	Sedang
4	Pendek	Kecil	Besar	Kecil
5	Pendek	Kecil	Kecil	Kecil
6	Pendek	Kecil	Besar	Kecil
7	Pendek	Kecil	Besar	Kecil
8	Pendek	Kecil	Besar	Kecil
9	Menengah	Kecil	Sedang	Sedang
10	Pendek	Kecil	Besar	Kecil
11	Panjang	Menengah	Sedang	Besar
12	Panjang	Menengah	Besar	Besar
13	Pendek	Kecil	Sedang	Kecil
14	Panjang	Menengah	Kecil	Besar
15	Menengah	Kecil	Besar	Sedang

Setelah itu dilakukan perhitungan dengan menggunakan algoritma C4.5 untuk mendapatkan decision tree. Perhitungan dimulai dengan menghitung node 1 dengan mencari nilai entropy dan nilai gain.

$$Entropy[Jumlah Pinjam, Pendek]=0$$

$$Entropy[Jumlah Pinjam, Menengah]=0$$

$$Entropy[Jumlah Pinjam, Panjang]=0.811278124$$

$$Entropy[Jangka Waktu, Kecil]=0.721928095$$

$$Entropy[Jangka Waktu, Menengah]=0.970950594$$

$$Entropy[Suku Bunga, Besar]=1.224394445$$

$$Entropy[Suku Bunga, Sedang]=1.584962501$$

$$Entropy[Suku Bunga, Kecil]=1.584962501$$

$$Total Entropy(S) = (-\frac{8}{15} * \log_2(\frac{8}{15})) + (-\frac{4}{15} * \log_2(\frac{4}{15})) + (-\frac{3}{15} * \log_2(\frac{3}{15})) = 1.456564762$$

$$Gain [ Total , Jumlah Pinjam ] = 1,456564762 - ((\frac{8}{15}*0)+(\frac{3}{15}*0)+(\frac{4}{15}*0,811278124))= 1.240223929$$

$$Gain [ Total , Jangka Waktu ] =$$

$$1,456564762 - ((\frac{10}{15}*0,721928095)+(\frac{5}{15}*0,970950594))= 0,651629167$$

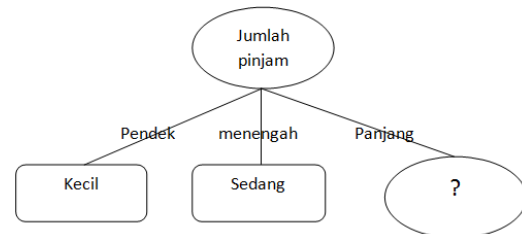
$$Gain [ Total , Suku Bunga ] = 1,456564762 - ((\frac{9}{15}*1,224394445)+(\frac{3}{15}*1,584962501)+(\frac{3}{15}*1,584962501)) = 0,651629167$$

Maka diperoleh hasil perhitungan pada node 1 seperti pada tabel dibawah ini

Tabel 3. Hasil Perhitungan pada Node 1

NODE		Jml Kasus (S)	Kecil (S <sub>1</sub> )	Sedang (S <sub>2</sub> )	Besar (S <sub>3</sub> )	Entropy	Gain
1	RESIKO	15	8	4	3	1,456564762	
	JUMLAH PINJAM						1,240231929
	PENDEK	8	8	0	0	0,000000000	
	MENENGAH	3	0	3	0	0,000000000	
	PANJANG	4	0	1	3	0,811248124	
	JANGKA WAKTU						0,651629167
	KECIL	10	8	2	0	0,721928095	
	MENENGAH	5	0	2	3	0,970950594	
	SUKU BUNGA						0,087943095
	BESAR	9	6	2	1	1,224394445	
	SEDANG	3	1	1	1	1,584962501	
	KECIL	3	1	1	1	1,584962501	

Dilihat dari tabel diatas, nilai gain tertinggi adalah Jumlah Pinjam, maka Jumlah Pinjam menjadi akar dari decision tree.



Gambar 2. Pohon Keputusan Pada Node 1

Dari decision tree pada node 1, dapat dilihat Jumlah Pinjam yang Panjang masih belum menjadi sebuah keputusan, maka dilakukan perhitungan untuk node 2 dengan menemukan entropy dan gain.

$$Total Entropy(S) = (-\frac{0}{4} * \log_2(\frac{0}{4})) + (-\frac{1}{4} * \log_2(\frac{1}{4})) + (-\frac{3}{4} * \log_2(\frac{3}{4})) = 0,811248124$$

$$Gain [ Total , Jangka Waktu ] = 0,811248124 - ((\frac{0}{4}*0)+(\frac{4}{4}*0,811248124)) = -2,188751876$$

$$Gain [ Total , Suku Bunga ] = 0,811248124 - ((\frac{2}{4}*1)+(\frac{1}{4}*0) + (\frac{1}{4}*0)) = 0,311248124$$

Maka diperoleh hasil perhitungan pada node 2 seperti pada tabel dibawah ini

Tabel 4. Hasil Perhitungan pada Node 2

NODE		Jml Kasus (S)	Kecil (S <sub>1</sub> )	Sedang (S <sub>2</sub> )	Besar (S <sub>3</sub> )	Entropy	Gain
2	JUMLAH PINJAM - PANJANG	4	0	1	3	0,811248124	-2,188751876
	JANGKA WAKTU						
	KECIL	0	0	0	0	0,000000000	
	MENENGAH	4	0	1	3	0,811278124	
	SUKU BUNGA						0,311248124
	BESAR	2	0	1	1	1,000000000	
	SEDANG	1	0	0	1	0,000000000	
	KECIL	1	0	0	1	0,000000000	

Dilihat dari tabel diatas, nilai gain tertinggi adalah Suku Bunga, maka Suku Bunga menjadi cabang selanjutnya dari decision tree.



Gambar 3. Pohon Keputusan Pada Node 2

Dari decision tree pada node 2, dapat dilihat Suku Bunga yang Besar masih belum menjadi sebuah keputusan, maka dilakukan perhitungan untuk node 3 dengan menemukan entropy dan gain.

$$Total\ Entropy(S) = (-\frac{0}{2} * \log_2(\frac{0}{2})) + (-\frac{1}{2} * \log_2(\frac{1}{2})) + (-\frac{1}{2} * \log_2(\frac{1}{2})) = 1,000000000$$

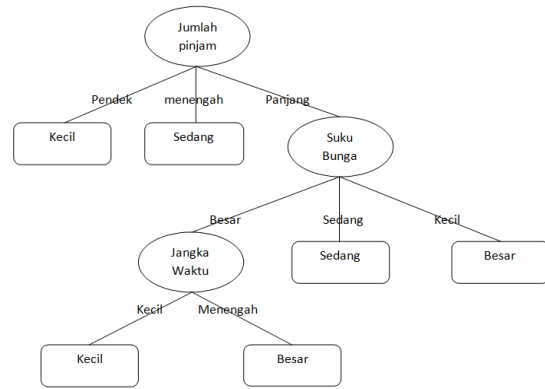
$$Gain [ Total, Jangka Waktu ] = 1,000000000 - ((\frac{0}{2} * 0) + (\frac{2}{2} * 1,000000000)) = 0,000000000$$

Maka diperoleh hasil perhitungan pada node 3 seperti pada tabel dibawah ini

Tabel 5. Hasil Perhitungan pada Node 3

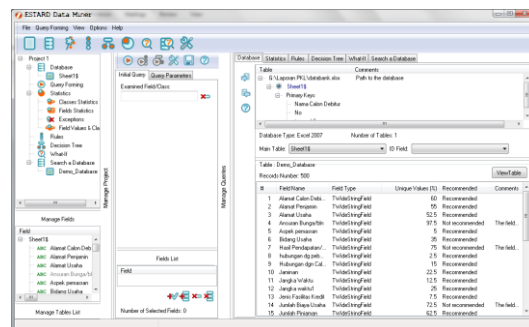
NODE		Jml Kasus (S)	Kecil (S <sub>1</sub> )	Sedang (S <sub>2</sub> )	Besar (S <sub>3</sub> )	Entropy	Gain
3	JUMLAH PINJAM - PANJANG dan SUKU BUNGA - BESAR	2	0	1	1	1,000000000	
	JANGKA WAKTU						0
	KECIL	0	0	0	0	0,000000000	
	MENENGAH	2	0	1	1	1,000000000	

Dilihat dari tabel diatas, dapat digambarkan decision tree seperti dibawah ini



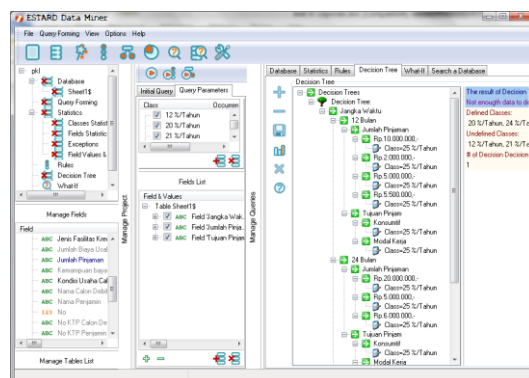
Gambar 4. Pohon Keputusan Pada Node 3

Hasil analisa akan diimplementasikan ke dalam software data mining yaitu ESTARD Data Miner. Data yang akan diolah dimasukkan ke dalam software ESTARD Data Miner seperti gambar berikut :



Gambar 5. Setelah Data Dimasukkan

Setelah data dimasukkan, akan diolah oleh software ESTARD Data Miner, sehingga akan menghasilkan decision tree seperti dibawah ini



Gambar 6. Decision Tree yang Dihasilkan ESTARD Data Miner

**KESIMPULAN**

Kesimpulan yang dapat dikemukakan :

1. Decision tree yang dihasilkan telah mampu menunjukkan besar atau kecilnya resiko

- kredit pada BPR (Bank Perkreditan Rakyat) Bukittandang Mandiri Padang.
2. Dengan adanya *Decision tree* ini ,dapat mempermudah pihak Bank dalam mengambil keputusan.
  3. Dengan meggunakan Estard Data Miner, hasil yang dapat dilihat berupa analisis keputusan yang menggambarkan presentase suku bunga terhadap kredit yang diberikan kepada nasabah dalam jangka waktu tertentu.
  4. Gambaran Presentase suku bunga tersebut tujuannya agar dapat mengetahui besar atau kecilnya resiko kredit. Maksudnya, apabila presentase suku bunga kredit dalam jangka terentu tinggi, maka resiko kredit yang dihadapi adalah kecil.

#### DAFTAR REFERENSI

- [1] Kusrini dan Emha Taufiq Luthfi. 2009. *Algoritma Data Mining*. Penerbit Andi Offset, Yogyakarta.
- [2] Santosa, Budi. 2007. *Data Mining : Teknik Pemanfaatan Data untuk keperluan Bisnis*. Graha Ilmu. Yogyakarta.
- [3] [http://id.wikipedia.org/w/index.php?title=Risiko\\_kredit&action=edit&section=4/](http://id.wikipedia.org/w/index.php?title=Risiko_kredit&action=edit&section=4/), diakses pada Juli 2011
- [4] <http://www.estard.com/>, diakses pada Juli 2011
- [5] Kusnawi. 2007. *Pengantar Solusi Data Mining*. Yogyakarta