

**PENERAPAN ALGORITMA KLASIFIKASI C4.5 UNTUK
PENENTUAN KELAYAKAN PEMBERIAN KREDIT
KOPERASI**



TESIS

FIRMANSYAH
14000216

PROGRAM PASCASARJANA MAGISTER ILMU KOMPUTER
SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
NUSA MANDIRI
JAKARTA
2011

**PENERAPAN ALGORITMA KLASIFIKASI C4.5 UNTUK
PENENTUAN KELAYAKAN PEMBERIAN KREDIT
KOPERASI**



TESIS

Diajukan sebagai salah satu syarat untuk memperoleh gelar
Magister Ilmu Komputer (M.Kom)

FIRMANSYAH
14000216

**PROGRAM PASCASARJANA MAGISTER ILMU KOMPUTER
SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
NUSA MANDIRI
JAKARTA
2011**

SURAT PERNYATAAN ORISINALITAS

Yang bertanda tangan di bawah ini :

Nama : Firmansyah
NIM : 14000216
Program Studi : Magister Ilmu Komputer
Jenjang : Strata Dua (S2)
Konsentrasi : *e-Business*

Dengan ini menyatakan bahwa tesis yang telah saya buat dengan judul: “Penerapan Algoritma Klasifikasi C4.5 untuk Penentuan Kelayakan Pemberian Kredit Koperasi” adalah hasil karya sendiri, dan semua sumber baik yang kutip maupun yang dirujuk telah saya nyatakan dengan benar dan tesis belum pernah diterbitkan atau dipublikasikan dimanapun dan dalam bentuk apapun.

Demikianlah surat pernyataan ini saya buat dengan sebenar-benarnya. Apabila dikemudian hari ternyata saya memberikan keterangan palsu dan atau ada pihak lain yang mengklaim bahwa tesis yang telah saya buat adalah hasil karya milik seseorang atau badan tertentu, saya bersedia diproses baik secara pidana maupun perdata dan kelulusan saya dari Program Pascasarjana Magister Ilmu Komputer Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri dicabut/dibatalkan.

Jakarta, 6 Februari 2011
Yang menyatakan,

Materai Rp. 6.000,-

Firmansyah

HALAMAN PENGESAHAN

Tesis ini diajukan oleh :

Nama : Firmansyah

NIM : 14000216

Program Studi : Magsiter Ilmu Komputer

Jenjang : Strata Dua (S2)

Konsentrasi : *e-Business*

Judul Tesis : “Penerapan Algoritma Klasifikasi C4.5 untuk Penentuan Kelayakan Pemberian Kredit Koperasi”

Telah berhasil dipertahankan dihadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Magister Ilmu Komputer (M.Kom) pada Program Pascasarjana Magister Ilmu Komputer Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri (STMIK Nusa Mandiri).

Jakarta, 24 April 2010
Pascasarjana Magister Ilmu Komputer
STMIK Nusa Mandiri
Direktur

H. Mochamad Wahyudi, MM, M.Kom

DEWAN PENGUJI

Penguji I : Dr. Dana Indra Sensuse

Penguji II : Drs. Bambang Setyarso, MA

Penguji III /
Pembimbing : Romi Satria Wahono, M.Eng

KATA PENGANTAR

Puji syukur alhamdulillah, penulis panjatkan kehadiran Allah, SWT, yang telah melimpahkan rahmat dan karunia-Nya, sehingga pada akhirnya penulis dapat menyelesaikan tesis ini tepat pada waktunya. Dimana tesis ini penulis sajikan dalam bentuk buku yang sederhana. Adapun judul tesis, yang penulis ambil sebagai berikut “Penerapan Algoritma Klasifikasi C4.5 untuk Penentuan Kelayakan Pemberian Kredit Koperasi”.

Tujuan penulisan tesis ini dibuat sebagai salah satu untuk mendapatkan gelar Magister Ilmu Komputer (M.Kom) pada Program Pascasarjana Magister Ilmu Komputer Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri (STMIK Nusa Mandiri).

Tesis ini diambil berdasarkan hasil penelitian yang penulis lakukan pada koperasi SUDIN Dikmenti Jakarta Selatan. Penulis juga lakukan mencari dan menganalisa berbagai macam sumber referensi, baik dalam bentuk jurnal ilmiah, buku-buku literatur, *internet*, dll yang terkait dengan pembahasan pada tesis ini.

Penulis menyadari bahwa tanpa bimbingan dan dukungan dari semua pihak dalam pembuatan tesis ini, maka penulis tidak dapat menyelesaikan tesis ini tepat pada waktunya. Untuk itu ijinkanlah penulis kesempatan ini untuk mengucapkan ucapan terima kasih yang sebesar-besarnya kepada :

1. Bapak Romi Satria Wahono, M.Eng selaku pembimbing tesis yang telah menyediakan waktu, pikiran dan tenaga dalam membimbing penulis dalam menyelesaikan tesis ini.
2. Bapak Andi Azis selaku Kepala Seksi SUDIN Dikmenti Jakarta Selatan yang telah mengijinkan penulis melakukan riset untuk mendapatkan data atau informasi yang penulis butuhkan.
3. Kedua orang tua yang telah memberikan dukungan material dan moral kepada penulis.
4. Seluruh staf pengajar (dosen) Program Pascasarjana Magister Ilmu Komputer Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri yang telah memberikan pelajaran yang berarti bagi penulis selama menempuh studi.

5. Seluruh staf dan karyawan Program Pascasarjana Magister Ilmu Komputer Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri yang telah melayani penulis dengan baik selama kuliah.
6. Para sahabat yang telah memberikan dukungan.
7. Rekan-rekan BTI BSI yang telah memberikan pengalamannya.
8. Semua yang tak bisa saya sebutkan satu persatu.

Serta semua pihak yang terlalu banyak untuk penulis sebutkan satu persatu sehingga terwujudnya penulisan tesis ini. Penulis menyadari bahwa penulisan tesis ini masih jauh sekali dari sempurna, untuk itu penulis mohon kritik dan saran yang bersifat membangun demi kesempurnaan penulisan karya ilmiah yang penulis hasilkan untuk yang akan datang.

Akhir kata semoga tesis ini dapat bermanfaat bagi penulis khususnya dan bagi para pembaca yang berminat pada umumnya.

Jakarta, 6 Februari 2011

Firmansyah

Penulis

**SURAT PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH
UNTUK KEPENTINGAN AKADEMIS**

Yang bertanda tangan di bawah ini, saya :

Nama : Firmansyah
NIM : 14000216
Program Studi : Magsiter Ilmu Komputer
Jenjang : Strata Dua (S2)
Konsentrasi : *e-Business*
Jenis Karya : Tesis

Demi pengembangan ilmu pengetahuan, dengan ini menyetujui untuk memberikan ijin kepada pihak Program Pascasarjana Magister Ilmu Komputer Sekolah Tinggi Manajemen Inbentukika dan Komputer Nusa Mandiri (STMIK Nusa Mandiri) **Hak Bebas Royalti Non-Eksklusif (*Non-exclusive Royalti-Free Right*)** atas karya ilmiah kami yang berjudul : “Penerapan Algoritma Klasifikasi C4.5 untuk Penentuan Kelayakan Pemberian Kredit Koperasi” beserta perangkat yang diperlukan (apabila ada).

Dengan **Hak Bebas Royalti Non-Eksklusif** ini pihak STMIK Nusa Mandiri berhak menyimpan, mengalih-media atau *bentuk*-kan, mengelolanya dalam pangkalan data (*database*), mendistribusikannya dan menampilkan atau mempublikasikannya di *internet* atau media lain untuk kepentingan akademis tanpa perlu meminta ijin dari kami selama tetap mencantumkan nama kami sebagai penulis/pencipta karya ilmiah tersebut.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak STMIK Nusa Mandiri, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 6 Februari 2011
Yang menyatakan,

Materai Rp. 6.000,-

Firmansyah

ABSTRAK

Nama : Firmansyah
NIM : 14000216
Program Studi : Magister Ilmu Komputer
Jenjang : Strata Dua (S2)
Konsentrasi : *e-Business*
Judul : “Penerapan Algoritma Klasifikasi C4.5 untuk Penentuan Kelayakan Pemberian Kredit Koperasi”

Kredit adalah penyediaan uang atau tagihan yang dapat dipersamakan dengan itu, berdasarkan persetujuan atau kesepakatan pinjam meminjam antara bank dengan pihak lain yang mewajibkan pihak peminjam melunasi utangnya setelah jangka waktu tertentu dengan pemberian bunga. Sebelum koperasi menyetujui kredit yang diajukan oleh debitur, koperasi melakukan analisis kredit terhadap debitur apakah pengajuan kredit disetujui atau tidak disetujui

Analisa kredit yang dilakukan oleh analis terkadang tidak akurat, sehingga beberapa debitur yang diberikan kredit tidak mempunyai kemampuan untuk membayar yang menyebabkan kredit macet. Dari permasalahan ini diperlukan suatu model yang mampu mengklasifikasikan sekaligus memprediksi debitur mana saja yang bermasalah dan tidak bermasalah.

Model pohon keputusan ternyata mampu meningkatkan akurasi dalam menganalisa kelayakan kredit yang diajukan calon debitur. Semakin kaya informasi atau pengetahuan yang dikandung oleh data training, maka akurasi pohon keputusan akan semakin meningkat.

Kata kunci:

Analisa kredit, pohon keputusan, algoritma C4.5

ABSTRACT

Name : Firmansyah
NIM : 14000216
Study of Program : Magister Ilmu Komputer
Levels : Strata Dua (S2)
Concentration : *e-Business*
Titel : “Penerapan Algoritma Klasifikasi C4.5 untuk Penentuan Kelayakan Pemberian Kredit Koperasi”

Credit is the provision of money or bills can be equivalent, under contracts lending between banks by another party that requires the borrower repay the debt after a certain period of time with interest. Prior to approve cooperative credit proposed by the debtor, the cooperative credit analysis of borrowers whether credit application is approved or not approved.

Credit analysis performed by analysts sometimes inaccurate, so some credit given debtor has no ability to pay that cause bad credit. From this problem we need a model that is able to classify at the same time predicting that troubled borrowers everywhere and not problematic.

Decision tree model was able to improve the accuracy in analyzing the credit worthiness of borrowers who filed. The more rich information or knowledge contained by the training data, the accuracy of decision trees will increase.

Keyword :

Credit analysis, decision tree, C4.5 algorithm

DAFTAR ISI

	Halaman
HALAMAN SAMPUL.....	i
HALAMAN JUDUL.....	ii
HALAMAN PERNYATAAN ORISINALITAS.....	iii
HALAMAN PENGESAHAN.....	iv
KATA PENGANTAR.....	v
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS.....	vii
ABSTRAK.....	viii
ABSTRACT.....	ix
DAFTAR ISI.....	x
DAFTAR TABEL.....	xi
DAFTAR GAMBAR.....	xii
DAFTAR LAMPIRAN.....	xiii
BAB 1. PENDAHULUAN.....	1
1.1. Latar Belakang Penulisan.....	1
1.2. Identifikasi Masalah.....	3
1.3. Tujuan Penelitian.....	3
1.4. Ruang Lingkup Penelitian.....	3
1.5. Manfaat penelitian.....	4
1.6. Ruang lingkup penelitian	4
1.7. Sistematika penulisan	4
BAB 2. LANDASAN/KERANGKA PEMIKIRAN.....	5
2.1. Tinjauan Studi.....	5
2.2. Tinjauan Pustaka.....	6
2.3. Studi kasus	14
2.4. Rapid Miner	19
2.5. Kerangka pemikiran	20
BAB 3. METODE PENELITIAN.....	22
BAB 4. HASIL PENELITIAN DAN PEMBAHASAN.....	42
BAB 5. PENUTUP.....	49
5.1. Kesimpulan.....	49
5.2. Saran.....	49
DAFTAR REFERENSI.....	50
SURAT KETERANGAN RISET/PRAKTEK KERJA LAPANGAN.....	52

DAFTAR TABEL

Tabel 1.1 Laporan data debitur bermasalah	2
Tabel 2.1 Atribut data debitur	7
Tabel 2.2 Tabel keputusan bermain tenis	14
Tabel 2.3 Perhitungan node 1	15
Tabel 2.4 Perhitungan node 1.1	16
Tabel 2.5 Perhitungan node 1.1.2	17
Tabel 3.1 Data kredit tahun 2003	23
Tabel 3.2 Data training kredit	24
Tabel 3.3 Kategori atribut	30
Tabel 3.4 Data debitur bermasalah	31
Tabel 3.5 Perhitungan node 1	32
Tabel 3.6 Perhitungan node 1.1 dan 1.2	33
Tabel 3.7 Perhitungan node 1.1.1 dan 1.2.1	33
Tabel 3.8 Perhitungan node 1.1.1.1 dan 1.1.1.2	34
Tabel 4.1 Data uji	42
Tabel 4.2 Hasil pengujian dengan C4.5	43
Tabel 4.3 Konversi ke <i>confusion matrix</i>	44
Tabel 4.4 Nilai <i>sensitivity, specificity, precision, dan accuracy</i> dalam persentase	45

DAFTAR GAMBAR

Gambar 1.1 Grafik peningkatan kredit macet	2
Gambar 2. 1 Data mining sebagai bagian dari proses <i>knowledge discovery</i>	8
Gambar 2. 2 (a) Tahap <i>learning</i> dan (b) tahap klasifikasi.	10
Gambar 2. 3 Contoh konsep pohon keputusan	11
Gambar 2. 4 Pohon Keputusan Hasil Perhitungan Node 1	16
Gambar 2. 5 Pohon keputusan hasil perhitungan Node 1.1	17
Gambar 2. 6 Pohon keputusan hasil perhitungan Node 1.1.2	18
Gambar 2. 7 Tampilan area kerja Rapid Miner	19
Gambar 2. 8 Tampilan hasil pohon keputusan	20
Gambar 2. 9 Kerangka pemikiran	21
Gambar 3. 1 Pohon keputusan dengan metode <i>information gain</i>	36
Gambar 3. 2 Tampilan langkah 1 : memilih data training	38
Gambar 3. 3 Tampilan langkah 4 : menentukan atribut dan label	39
Gambar 3. 4 Relasi antara data training dan model	40
Gambar 3.5 Pohon keputusan yang telah terbentuk	40
Gambar 4. 1 Model <i>confusion matrix</i>	44
Gambar 4.2 Confusion matrix dengan RapidMiner	46
Gambar 4.3 Nilai AUC dalam grafik ROC dengan metode <i>information gain</i> dan <i>gain ratio</i> dengan memasukkan data uji.....	46
Gambar 4.4 Nilai AUC dalam grafik ROC metode <i>gini index</i> dengan memasukkan data uji	47
Gambar 4.5 Grafik ROC dengan perbandingan tiga metode, yaitu <i>information gain</i> , <i>gain ratio</i> dan <i>gini index</i> tanpa memassukan data uji	47

DAFTAR LAMPIRAN

Lampiran 1. Laporan kredit tahun 2003	53
---	----

BAB I

PENDAHULUAN

2.1 Latar Belakang Penulisan

Kredit adalah penyediaan uang atau tagihan yang dapat dipersamakan dengan itu, berdasarkan persetujuan atau kesepakatan pinjam meminjam antara bank dengan pihak lain yang mewajibkan pihak peminjam melunasi utangnya setelah jangka waktu tertentu dengan pemberian bunga. Sedangkan Kredit Tanpa Jaminan atau Kredit Tanpa Agunan (KTA) adalah kredit yang diberikan bank atau lembaga keuangan dalam bentuk uang tunai, yang dapat diperoleh tanpa memberikan jaminan. Kemudahan ini banyak dimanfaatkan oleh para nasabah untuk meminjam sejumlah dana dari koperasi untuk digunakan berbagai keperluan nasabah (UU Perbankan No. 10 tahun 1998).

Koperasi adalah merupakan usaha kekeluargaan dengan tujuan untuk mensejahterakan anggotanya (UUD 1945 pasal 33 ayat 1). Umumnya koperasi dikendalikan secara bersama oleh seluruh anggotanya dimana setiap anggota memiliki hak suara yang sama dalam setiap keputusan yang diambil. Ada beberapa tahap yang harus dilalui sebelum pinjaman dicairkan, mulai dari calon debitur mengisi data dan mengajukan aplikasi, setelah aplikasi diterima oleh koperasi, maka koperasi akan melakukan survey untuk memastikan apakah data yang diisi oleh calon debitur benar adanya. Apabila tidak ada masalah dalam proses survey, koperasi akan melakukan analisa berdasarkan aplikasi calon debitur, termasuk latar belakang, kemampuan debitur dan lain-lain. Jika tahap ini sudah disetujui maka dana yang diajukan oleh debitur akan dicairkan.

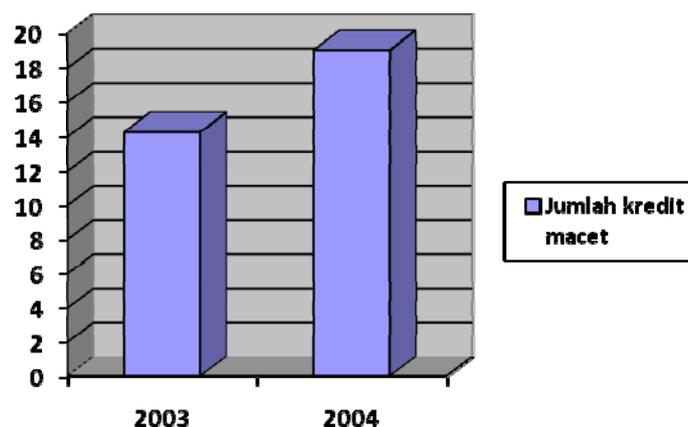
Kredit yang diajukan oleh debitur memiliki resiko, karena dari sekian banyak debitur yang mengajukan aplikasi ada kemungkinan beberapa debitur yang bermasalah dalam pembayarannya sehingga menyebabkan kredit macet. Sebelum koperasi menyetujui kredit yang diajukan oleh debitur, koperasi melakukan analisis kredit terhadap debitur apakah pengajuan kredit disetujui atau tidak disetujui. Analisa kredit adalah penyelidikan faktor-faktor yang berpengaruh pada lancarnya atau kurang lancarnya pengembalian kredit (Basuki, 2007).

Analisis kredit merupakan hal yang penting dalam lingkup resiko keuangan (Lai, Yu, Zhou, & Wang, 2006), oleh karena itu perlunya dilakukannya analisa. Namun, melakukan proses analisa kredit membutuhkan waktu lama (Kotsiantis, Kanellopoulos, Karioti, & Tampakas, 2009) dan mengidentifikasi data debitur yang bermasalah merupakan hal yang sulit (Odeh, Featherstone, & Das, 2010). Data yang besar dan banyaknya parameter tentunya membutuhkan alat yang efektif dan efisien untuk melakukan analisa kredit dan menilai debitur yang mempunyai resiko gagal bayar dan yang tidak beresiko.

Sebagai tolak ukur bahwa debitur disetujui atau ditolak, dapat digunakan data histori debitur yang telah disetujui oleh koperasi. Namun, perlu diperhatikan juga bahwa debitur yang telah disetujui juga tidak semuanya pembayar kredit yang baik, artinya ada beberapa debitur yang telah disetujui tapi beberapa bulan kemudian pembayarannya menunggak. Dibawah ini merupakan tabel yang menunjukkan data debitur bermasalah disebabkan ketidakakuratan analisis dalam menganalisa kredit yang diajukan debitur sehingga menyebabkan kredit macet, data ini diambil dari koperasi SUDIN Pendidikan Jakarta Selatan yaitu :

Tabel 1.1 Laporan Data Debitur Bermasalah dari tahun 2003 s.d 2004 dalam persentase (sumber : laporan koperasi Dikmenti Jaksel 2003-2004)

Tahun	Jumlah Anggota	Kredit Macet
2003	239	12.13%
2004	259	12.74%



Gambar 2.1 Grafik peningkatan kredit macet dari tahun 2003 s.d 2004 (sumber : laporan simpan pinjam koperasi SUDIN Dikmenti Jaksel)

Tabel 1.1 menunjukkan mulai tahun 2003 hingga tahun 2004 terjadi kenaikan kredit macet, ini disebabkan karena sejumlah anggota koperasi yang menunggak pembayaran kredit. Data di atas menjadi permasalahan bagi koperasi, karena jika debitur tidak mampu membayar cicilan kredit dalam jangka waktu yang telah ditentukan, maka ini akan menyebabkan kredit macet.

Ada beberapa penelitian dan tehnik analisa kredit yang dibuat oleh beberapa peneliti seperti Lai, Yu, Zhou dan Wang (2006) yang melakukan evaluasi resiko kredit dengan menggunakan model Least Square Support Vector Machines. Henley dan Hand (1996) juga melakukan penilaian terhadap kredit debitur dengan menggunakan model K Nearest Neighbour. Jiang (2009) membuat model untuk memprediksi nasabah yang yang bermasalah dan tidak bermasalah dalam pembayaran kredit dengan menggunakan model algoritma C4.5. Data yang digunakan diambil dari perusahaan German credit yang merupakan perusahaan pembiayaan. Jiang mengambil beberapa atribut dan kemudian dimasukkan ke dalam model untuk memprediksi persentase nasabah yang bermasalah.

Untuk mengatasi permasalahan di atas, maka penulis menggunakan model pohon keputusan algoritma C4.5. Model ini akan digunakan untuk memprediksi apakah debitur nantinya akan bermasalah dalam pembayaran kredit atau tidak. Model ini dipilih karena proses learning dan klasifikasi pada algoritma C4.5 sederhana dan cepat. Secara umum, model algoritma C4.5 mempunyai tingkat akurasi yang tinggi (Han & Kamber, 2006).

2.2 Identifikasi Masalah

Permasalahan yang muncul berdasarkan uraian di atas adalah banyaknya hasil analisis yang tidak akurat, seperti pengajuan kredit yang disetujui ternyata bermasalah dalam pembayaran cicilan.

2.3 Tujuan Penelitian

Dengan menerapkan Pohon Keputusan berbasis algoritma C4.5 diharapkan dapat meningkatkan keakuratan analisa kredit.

2.4 Manfaat Penelitian

Manfaat dari penelitian ini adalah :

- a. Manfaat praktis dari penelitian ini adalah untuk memudahkan analisis kredit dalam memberikan keputusan.
- b. Manfaat teoritis penelitian ini yaitu untuk memberikan sumbangan penerapan model algoritma C4.5.
- c. Manfaat kebijakan yaitu agar algoritma C4.5 mampu menjadi alat pendukung keputusan analisis kredit.

2.5 Ruang Lingkup Penelitian

Ruang lingkup berfungsi untuk membatasi pembahasan pada pokok permasalahan saja. Ruang lingkup wilayah penelitian yaitu mencakup Koperasi SUDIN Jakarta Selatan, sedangkan ruang lingkup waktu yaitu dimulainya penelitian hingga selesainya observasi yang dilakukan yaitu tanggal 1 Desember 2010 s.d 10 Januari 2011. Data yang diteliti mulai tahun 2003 hingga tahun 2004.

2.6 Sistematika Penulisan

Bab I : Pendahuluan. Membahas mengenai latar belakang penulisan, masalah kredit yang terjadi di koperasi, pemecahan masalah dan tujuan penelitian.

Bab II : Landasan teori. Membahas tentang teori yang melandasi penelitian yaitu model pohon keputusan dan algoritma C4.5. Studi kasus disajikan untuk memberi contoh dan langkah algoritma C4.5.

Bab III : Metode penelitian. Membahas metode pengumpulan data dan eksperimen. Eksperimen merupakan inti pembahasan dari bab ini, yaitu menguji model algoritma C4.5 untuk memprediksi debitur bermasalah.

BAB IV : Hasil dan pembahasan. Menampilkan hasil dari eksperimen, baik sebelum diterapkan model maupun setelah diterapkan model. Membahas metode mengukur hasil dengan metode statistik. Ditampilkan hasil uji dalam model statistik.

BAB V : Penutup. Membahas kesimpulan dari penelitian dan saran untuk penelitian selanjutnya.

BAB II

LANDASAN/KERANGKA PEMIKIRAN

2.1 Tinjauan Studi

Ada beberapa penelitian yang menggunakan model Pohon Keputusan C4.5 sebagai algoritma untuk memprediksi berdasarkan data histori.

1. Credit Scoring Model Based on Decision Tree and the Simulated Annealing Algorithm (Jiang, 2006).

Membuat model untuk memprediksi nasabah yang bermasalah dan tidak bermasalah dalam pembayaran kredit dengan menggunakan model Pohon Keputusan dan C4.5 dan Simulated Annealing Algoritma. Data yang digunakan diambil dari perusahaan German credit yang merupakan perusahaan pembiayaan. Jiang mengambil beberapa atribut dan kemudian dimasukkan ke dalam model untuk memprediksi persentase nasabah yang bermasalah.

2. Could Decision Trees Improve the Classification Accuracy and Interpretability of Loan Granting Decisions ? (Zurada, 2010).

Membuat penelitian untuk membandingkan beberapa algoritma seperti Regresi Linier, Neural Network, Support Vector Machine, Case Base Reasoning, Rule Based Fuzzy Neural Network dan Decision Tree. Semua model algoritma tadi digunakan untuk menganalisa persetujuan pinjaman dalam bentuk kredit. Dari hasil penelitian didapatkan bahwa Decision Tree terbukti mempunyai akurasi tertinggi dalam menentukan keputusan dibandingkan algoritma lain.

3. Comparing decision trees with logistic regression for credit risk analysis (Satchidananda & Simha, 2006).

Penelitian ini membandingkan dua model algoritma untuk analisa resiko kredit, yaitu Pohon Keputusan dan Regresi Logistik. Data diambil dari dua bank yang berbeda, kemudian untuk mengelompokkan kasus positif dan negatif maka dilakukan klustering data dengan menggunakan k-means. Hasil analisa dari masing-masing model dikomparasi dan kemudian diukur.,

kemudian didapatkan bahwa algoritma pohon keputusan mempunyai tingkat akurasi yang tinggi dibandingkan algoritma regresi logistik.

Dari semua penelitian dan model yang digunakan, terbukti bahwa pohon keputusan dan algoritma C4.5 mampu menyelesaikan masalah yang berhubungan dengan analisa kredit dan membuktikan pohon keputusan menghasilkan akurasi yang lebih baik dibanding model lain.

2.2 Tinjauan Pustaka

Dalam penulisan tesis ini, penulis menggunakan buku, prosiding, dan jurnal sebagai referensi untuk menjelaskan model algoritma C4.5.

2.2.1 Kredit

Pengertian kredit pada pasal 1 angka 11 Undang-undang no 10 tahun 1998 tentang perubahan undang-undang no 7 tahun 1992 tentang Perbankan adalah penyediaan uang atau tagihan yang dapat dipersamakan dengan itu, berdasarkan persetujuan atau kesepakatan pinjam meminjam antar bank dengan pihak lain yang mewajibkan pihak peminjam untuk melunasi utangnya dalam jangka waktu tertentu dengan pemberian bunga.

Ada beberapa prinsip penilaian kredit yang sering dilakukan, yaitu prinsip 5C (Kasmir, 2000) dalam (Triwahyuniati, 2008):

1. *Character*, adalah watak atau sifat seseorang dalam hal ini adalah calon debitur. Tujuannya adalah untuk memberikan keyakinan kepada bank, bahwa watak atau sifat dari orang-orang yang akan diberikan kredit benar-benar dapat dipercaya.
2. *Capacity*, untuk melihat kemampuan calon nasabah dalam membayar kredit dihubungkan dengan kemampuan mengelola bisnis serta kemampuan mencari laba.
3. *Capital*, dimana dimana untuk mengetahui sumber-sumber pembiayaan yang dimiliki nasabah terhadap usaha yang akan dibiayai oleh bank.
4. *Collateral*, merupakan jaminan yang diberikan calon nasabah baik yang bersifat fisik maupun non fisik.
5. *Condition*, dalam menilai kredit hendaknya dinilai kondisi ekonomi sekarang dan untuk di masa yang akan datang sesuai sektor masing-masing.

Tehnik analisa kredit dengan menggunakan algoritma C4.5 diharapkan dapat meminimalkan masuknya debitur yang bermasalah, karena semakin banyak debitur bermasalah maka akan semakin memperbesar tingkat kredit macet yang pada akhirnya dapat menimbulkan kebangkrutan. Ditulis dalam (Zurada, 2010) bahwa ada beberapa atribut yang menyertai data debitur yaitu seperti umur, jumlah kredit, checking, penjamin, jangka kredit, lama bekerja, jumlah akun di bank, status pekerjaan, histori kredit, status rumah, dana aman, status pernikahan, alasan pinjaman, dan lain-lain. Namun dalam penelitian ini hanya beberapa atribut penting yang digunakan sebagai ukuran untuk menganalisa calon debitur, yaitu :

Tabel 2. 1 Atribut data debitur dibagi berdasarkan nilai kategorikal dan nilai kontinu.

Nama Atribut	Sifat Nilai
Jumlah Tanggungan	Kategorikal
Level Pekerjaan	Kategorikal
Jumlah Pinjaman	Nilai Kontinu
Jangka Waktu	Kategorikal
Penghasilan	Nilai Kontinu
Class	Kategorikal

Pada tabel 2.1 di atas sifat nilai dibagi dua yaitu kategorikal dan nilai kontinu, kategorikal merupakan nilai yang bukan dalam bentuk angka seperti jumlah tanggungan dibagi berdasarkan kategori sedikit, sedang dan banyak. Sedangkan nilai kontinu adalah nilai atribut yang bernilai angka, seperti jumlah pinjaman Rp. 10.000.000, merupakan nilai nominal. Sebenarnya nilai kategorikal merupakan nilai yang diambil dari nilai angka, namun untuk memudahkan dalam proses analisa maka angka-angka tersebut dibagi menjadi beberapa kategori.

2.2.2 Data Mining

Data mining merupakan disebut sebagai proses ekstraksi pengetahuan dari data yang besar. Sesuai fungsinya, data mining adalah proses pengambilan pengetahuan dari volume data yang besar yang disimpan dalam basis data, data *warehouse*, atau informasi yang disimpan dalam repositori (Han & Kamber, 2006). Gartner Group dalam (Larose, 2005) menyebutkan bahwa data mining adalah proses menelusuri pengetahuan baru, pola dan tren yang dipilah dari jumlah data yang besar yang disimpan dalam repositori atau tempat penyimpanan