

KLASIFIKASI DATA MINING UNTUK MENENTUKAN TINGKAT PERSETUJUAN KARTU KREDIT

Oleh :

Hanggi Dwifa Honesqi

Program Studi Sistem Informasi
Universitas Putra Indonesia – YPTK
Padang, Indonesia
hanggihonesqi@gmail.com

ABSTRAK

Algoritma C4.5 adalah algoritma klasifikasi data bertipe pohon keputusan (Decision Tree). Pohon keputusan algoritma C4.5 dibangun dengan tiga tahap yaitu pemilihan atribut sebagai akar, membuat cabang untuk tiap-tiap nilai dan membagi kasus dalam cabang. Tahap-tahap ini akan diulangi untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama. Dari penyelesaian pohon keputusan maka akan di hasilkan beberapa *Rule* atau pengetahuan dari suatu kasus. Dalam tesis ini penulis mengklasifikasikan kriteria calon nasabah yang membuat kartu kredit di PT. Cipta Daya *Resoure inhouse* BNI46. Pengklasifikasian nasabah ini sangat di tentukan oleh variabel-variabel input yang di tetapkan antara lain jenis perusahaan, status pekerjaan, jabatan, lama kerja, gaji perbulan, tanggal lahir dan status tempat tinggal. Dengan penerapan algoritma C4.5 ini akan dapat membantu pihak Manajemen Marketing PT. Cipta Daya *Resoure inhouse* BNI46 dalam menentukan kriteria calon nasabah yang tepat untuk memperoleh kartu kredit.

Kata kunci : Data Mining, klasifikasi, algoritma C4.5, pohon keputusan, Kartu kredit, nasabah.

1. PENDAHULUAN

Sistem kartu kredit adalah suatu jenis penyelesaian transaksi ritel (retail) dan sistem kredit, yang namanya berasal dari kartu plastik yang diterbitkan kepada pengguna sistem tersebut. Sebuah kartu kredit berbeda dengan kartu debit di mana penerbit kartu kredit meminjamkan konsumen uang dan bukan mengambil uang dari rekening. Di dalam proses permohonan dan penerbitan kartu kredit ada beberapa tahapan yang harus dilalui, Persyaratan-persyaratan tersebut adalah Mengisi Data pribadi, Data pekerjaan, Data penghasilan dan referensi Bank, Data lainnya yang dimaksud dengan pekerjaan dapat berwiraswasta atau pegawai swasta atau kalangan profesional tertentu.

Data mining adalah proses menganalisa data dari perspektif yang berbeda dan menyimpulkannya menjadi informasi-informasi penting yang dapat dipakai untuk meningkatkan keuntungan, memperkecil biaya

pengeluaran, atau bahkan keduanya (Angga & Riani,2012). Disini dapat disimpulkan bahwa data mining berfungsi salah satunya untuk memperkecil biaya pengeluaran. Dalam analisis kartu kredit apabila data telah lengkap semuanya sesuai dengan data diatas maka akan dilakukan cek data via telepon untuk konfirmasi data kepada nasabah apakah data sesuai dengan kebenarannya. Dalam hal ini banyak terjadi kesalahan pada analisis dikarenakan banyaknya permintaan kartu kredit sehingga analisis harus menelepon data nasabah satu persatu untuk memastikan bahwa data yang diajukan ini adalah benar. Dengan demikian butuh waktu lama untuk memprosesnya agar kartu kredit itu disetujui.

Pohon Keputusan adalah sebuah struktur yang dapat digunakan untuk membagi kumpulan data yang besar menjadi himpunan-himpunan *record* yang lebih kecil dengan menerapkan serangkaian aturan keputusan (Berry & Linoff,2004) & (Kusrini & Luthfi,

2009). Dengan menghimpun *record* yang lebih kecil dengan serangkaian aturan keputusan. Analisis tidak perlu menelepon semua nasabah yang mengajukan aplikasi kartu kredit. Dengan menentukan tingkat persetujuan kartu kredit analisis akan diberikan keputusan untuk memproses dan menolak aplikasi kartu kredit nasabah.

Oleh karenanya, penelitian ini dilakukan untuk membantu menyelesaikan permasalahan tersebut dengan merancang sebuah aplikasi data mining yang berfungsi untuk menentukan tingkat persetujuan kartu kredit dengan begitu analisis dapat mengurangi tingkat biaya via telepon dalam memproses persetujuan kartu kredit. Teknik data mining yang diterapkan pada aplikasi yang dibangun adalah Algoritma C4.5 atau sering disebut pohon keputusan merupakan metode klasifikasi dan prediksi yang sangat kuat dan terkenal. Metode tersebut bisa digunakan salah satunya untuk tujuan mempermudah menganalisa calon nasabah kartu kredit, dikarenakan mulai banyaknya data fiktif yang masuk dan juga menghindarkan kredit macet yang terjadi pada kartu kredit. Dengan pemilihan nasabah yang tepat dan sesuai pasar akan membuat keuntungan bank menjadi lebih baik. Diketahui dari infobank bahwa kartu kredit mengalami kenaikan dalam kemacetan pada tahun 2013 sebesar 2.4%. Untuk itu dengan ada sistem data mining untuk menentukan tingkat persetujuan kartu kredit dapat membantu dalam menentukan dan menganalisa persetujuan kartu kredit. Tujuannya adalah agar kita dapat mengetahui tindak lanjut berikutnya yang dapat diambil. Semua hal tersebut bertujuan mendukung kegiatan operasional perusahaan sehingga tujuan akhir perusahaan diharapkan dapat tercapai. Salah satu contoh klasifikasi data mining untuk menentukan tingkat persetujuan kartu kredit dengan menggunakan algoritma C4.5.

2. METODE PENELITIAN

1. Pengumpulan Data

Data sekunder pada penelitian ini adalah : buku-buku, jurnal tentang pohon keputusan algoritma C4.5 dan data mining serta data pengusul kartu kredit pada PT. Cipta Daya *Resoure* anak perusahaan BNI46. Sedangkan Data primer adalah data yang diperoleh dari hasil penelitian. Data primer dalam penelitian ini adalah data hasil uji dengan menggunakan algoritma Decision Tree C4.5

2. Pengolahan Awal Data

Proses pengolahan awal data dilakukan untuk mendapatkan fakta yang dapat digunakan untuk memecahkan masalah yang sedang diteliti. Dibatasi dalam menganalisa data berupa jenis perusahaan, status pekerjaan, jabatan, masa kerja, gaji perbulan dan tempat tinggal. Analisa ini digunakan untuk melihat masalah yang sedang terjadi yang akan digunakan sebagai bahan yang akan diteliti. Metode yang digunakan dalam penelitian ini adalah data *mining* dengan menggunakan algoritma 4.5. Di mana algoritma 4.5 digunakan sebagai teknik pembelajaran pada data mining tersebut.

3. Model Perancangan

Model yang diusulkan untuk menentukan tingkat persetujuan kartu kredit adalah dengan menggunakan pohon keputusan algoritma C4.5. Langkah-langkah untuk membangun sebuah pohon keputusan pada algoritma C4.5 adalah sebagai berikut (Kusrini & Luthfi,2009) :

- a. Pilih atribut sebagai akar
- b. Buat cabang untuk tiap-tiap nilai
- c. Bagi kasus dengan cabang

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

Untuk memilih atribut sebagai akar, didasarkan pada nilai *gain* tertinggi dari atribut-atribut yang ada. Untuk menghitung *gain* dapat digunakan rumus sebagai berikut (Kusrini & Luthfi, 2009) :

$$Gain(S,A) = Entropy(S) - \sum_{i=1}^n \frac{s_i}{s} + Entropy(s_i)$$

Di mana :

S : Himpunan kasus

A : atribut

n : jumlah partisi atribut A

$|S_i|$: jumlah kasus pada partisi ke- i

$|S|$: jumlah kasus pada S

Untuk menghitung nilai *entropy* dapat digunakan rumus sebagai berikut (Kusrini & Luthfi, 2009) :

$$Entropy(s) = \sum_{i=1}^n -P_i * \log_2 P_i \dots\dots\dots$$

Di mana:

S : Himpunan kasus

A : fitur

n : jumlah partisi S

P_i : proporsi dari S_i terhadap S

4. Implementasi Sistem

Pada tahapan implementasi bertujuan untuk mendapatkan hasil dari analisa keputusan terhadap calon nasabah. Setelah data dianalisa ,kemudian dilakukan pelatihan untuk mendapatkan hasil yang diinginkan. Pelatihan dilakukan dengan menggunakan dua tahapan. Tahapan – tahapan yang digunakan dalam pelatihan yaitu :

- a. Pelatihan secara manual

Pelatihan secara manual dilakukan dengan mengolah data yang telah didapatkan dari proses analisa manual.

Kemudian data tersebut dimasukkan ke dalam rumus – rumus perhitungan algoritma C4.5.

- b. Pelatihan secara *software*

Data yang telah dianalisa akan dilakukan pelatihan secara *software*. Pelatihan secara *software* dilakukan dengan menggunakan *software Rapid Miner*. Hal ini sangat berpengaruh terhadap penelitian, oleh karena itu data tersebut harus akurat agar dapat digunakan untuk meningkatkan kualitas layanan pada PT Cipta Daya *Resoure* anak perusahaan BNI46.

5. Pengujian Hasil

Pada tahapan pengujian dilakukan untuk mengetahui cara sistem bekerja dalam mengolah data. Data yang diperoleh berupa data nasabah yang mengajukan kartu kredit pada bulan januari 2014 yang akan telah diperoleh data 101 nasabah dan akan di uji 40% dari data yang ada. Poin-poin yang akan diambil dari jenis perusahaan, status pekerjaan, jabatan, masa kerja, gaji perbulan, dan tempat tinggal. atribut tersebut bisa mendapatkan hasil yang diinginkan dengan data yang telah melalui pelatihan selanjutnya akan dilakukan pengujian secara *software*. Dalam hasil pengujian ini akan didapatkan hasil keputusan diproses dan ditolak. Pengujian secara *software* dilakukan dengan menggunakan aplikasi *Rapid Miner*.

3. ANALISA DAN PERANCANGAN

Proses klasifikasi kelayakan calon nasabah baru diproses dengan membuat pohon keputusan untuk menghasilkan suatu *output*. Di mana data tersebut mempunyai beberapa kriteria yang merupakan syarat dalam pengklasifikasian kelulusan kelayakan calon

nasabah baru. Data tersebut merupakan data *input* klasifikasi kelayakan calon nasabah baru. Data yang telah didapat dianalisa sehingga dikelompokkan menjadi beberapa kelompok data yang dapat diproses dengan merancang pohon keputusan. Proses pengambilan keputusan dalam klasifikasi kelayakan calon nasabah baru dikelompokkan menjadi beberapa kriteria yaitu; jenis perusahaan, status pekerjaan, jabatan, lama kerja/usaha, gaji perbulan, usia, dan status tempat tinggal.

Dari data-data tersebut, yang diambil sebagai variabel keputusannya adalah komite diproses dan ditolak. Sedangkan yang diambil sebagai variabel penentu dalam pembentukan pohon keputusan adalah Jenis Perusahaan, Status Pekerjaan, Jabatan, Lama Kerja, Gaji Perbulan, Tanggal lahir dan Status Tempat Tinggal. Pemilihan variabel-variabel tersebut dengan pertimbangan bahwa jumlah nilai variabelnya tidak banyak sehingga diharapkan calon nasabah yang masuk dalam satu klasifikasi nilai variabel tersebut cukup banyak.

Dari *format* data akhir calon nasabah baru maka akan dilakukan klasifikasi data algoritma C4.5 dengan membuat pohon keputusan. Untuk memilih atribut sebagai akar, didasarkan pada nilai *gain* tertinggi dari atribut-atribut yang ada. Untuk menghitung *gain* digunakan rumus (1), sedangkan untuk menghitung nilai *entropy* dapat dilihat pada rumus (2). Setelah nilai *Entropy* dan *Gain* dihitung. Dan dari perhitungan dapat diketahui bahwa atribut dengan *gain* tertinggi adalah Jabatan sebesar **0,71127**. Berarti Jabatan dapat menjadi *node* akar. Kemudian, langkah selanjutnya adalah menyelesaikan untuk menghitung *Node* 1 sebagai akar, sama dengan cara yang di atas dengan menghitung nilai *entropy* dari atribut yang tersisa yaitu Gaji Perbulan, Usia, Lama Kerja, Jenis Perusahaan, Status Tempat Tinggal dan Status Pekerjaan setelah dihitung *entropy*, kemudian menghitung *gain* untuk tiap-tiap atribut.

Dari pohon keputusan yang didapat bahwa secara umum calon nasabah di PT. Cipta Daya *Resoure inhouse* BNI46

mempunyai Jabatan Customer Service, Dokter, Dosen dan Dokter Bedah, Guru, Guru kelas, Kepala Sekolah, Keuangan, Koordinator, Manager, Manager Marketing, Manager Pemasaran, Pemilik, Perawat, Pimpinan Cabang, Sales Executive, Senior Marketing, Staff, Supervisor Internal dan Lama Kerja dengan klasifikasi 2 sampai 13 dan Usia Nasabah dengan klasifikasi 1 sampai 20.

Berdasarkan pohon keputusan terakhir yang terbentuk, maka aturan atau *rule* yang terbentuk adalah sebagai berikut :

1. *If* Jabatan = Administrasi *Then* keputusan = Tolak
2. *If* Jabatan = Crew Counter *Then* keputusan = Tolak
3. *If* Jabatan = Customer Service and Lama Kerja = 1 *Then* Keputusan = 1
4. *If* Jabatan = Customer Service and Lama Kerja = 1 *Then* Keputusan Tolak
5. *If* Jabatan = Customer Service and Lama Kerja = 2 *Then* Keputusan = Diproses
6. *If* Jabatan = Dokter *Then* keputusan = Proses
7. *If* Jabatan = Dosen Dan Dokter Bedah *Then* keputusan = Proses
8. *If* Jabatan = Front Officer *Then* keputusan = Tolak
9. *If* Jabatan = Guru *Then* keputusan = Proses
10. *If* Jabatan = Guru Kelas *Then* keputusan = Proses
11. *If* Jabatan = Kepala Sekolah *Then* keputusan = Proses
12. *If* Jabatan = Keuangan *Then* keputusan = Proses
13. *If* Jabatan = Koordinator *Then* keputusan = Proses
14. *If* Jabatan = Manager *Then* keputusan = Proses
15. *If* Jabatan = Manager Makerting *Then* keputusan = Proses
16. *If* Jabatan = Manager Pemasaran *Then* keputusan = Proses
17. *If* Jabatan = Operasional Officer *Then* keputusan = Tolak

18. *If* Jabatan = Operational *Then* keputusan = Tolak
19. *If* Jabatan = Pemilik and Usia Nasabah = 2 *Then* Keputusan = 2
20. *If* Jabatan = Pemilik and Usia Nasabah = 21 *Then* Keputusan Tolak
21. *If* Jabatan = Pemilik and Usia Nasabah = 11 *Then* Keputusan = Diproses
22. *If* Jabatan = Perawat *Then* keputusan = Proses
23. *If* Jabatan = Petugas *Then* keputusan = Tolak
24. *If* Jabatan = Pimpinan Cabang *Then* keputusan = Proses
25. *If* Jabatan = Sales Counter *Then* keputusan = Tolak
26. *If* Jabatan = Sales Eksekutif *Then* keputusan = Proses
27. *If* Jabatan = Sekretaris Eksekutif *Then* keputusan = Tolak
28. *If* Jabatan = Senior Makerting *Then* keputusan = Proses
29. *If* Jabatan = Staff *Then* keputusan = Proses
30. *If* Jabatan = Supervisor Internal *Then* keputusan = Proses

Sebagai contoh hasil akhir, seorang calon nasabah dengan id (Lolynda) adalah seorang calon nasabah yang bekerja di perusahaan BUMD yaitu BANK NAGARI SOLOK dengan jabatan sebagai customer service dengan status pekerjaan karyawan yang sudah bekerja selama 3 Tahun. Apabila semuanya telah memenuhi syarat maka calon nasabah tersebut akan diproses dan apabila calon nasabah tersebut tidak memenuhi kriteria yang ditetapkan oleh pihak bank maka akan ditolak.

4. KESIMPULAN

Dari hasil uji yang dilakukan, metode pohon keputusan (*decision tree*) yang diproses dengan *software Rapidminer* lebih efektif dan fleksibel jika digunakan pada proses pengklasifikasian calon nasabah pada PT.

Cipta Daya *Resoure inhouse* BNI46. Disamping itu, pemilihan variabel (atribut kondisi dan attribut keputusan) yang akan digunakan dalam menentukan sebuah klasifikasi juga sangat mempengaruhi *rule* atau *knowledge* yang dihasilkan. Dalam hal ini Algoritma *C4.5* dianggap sebagai algoritma yang sangat membantu dalam melakukan klasifikasi data karena karakteristik data yang diklasifikasi dapat diperoleh dengan jelas, baik dalam bentuk struktur pohon keputusan (*decision tree*) maupun dalam aturan rule *If – Then* sehingga memudahkan pengguna dalam melakukan penggalian informasi terhadap data yang bersangkutan.

DAFTAR PUSTAKA

- Mabrur, Angga Ginanjar dan Riani Lubis. 2012. “Penerapan Data Mining Untuk Memprediksi Kriteria Nasabah Kredit”. *Jurnal Komputer dan Informatika (KOMPUTA)*.
- Kusriani dan Emha Taufiq Luthfi. 2009. *Algoritma Data Mining*. Yogyakarta: Andi.
- Widodo, Prabowo Pudjo. Dkk. 2013. Penerapan Data Mining Dengan Matlab. Bandung : Rekayasa Sains.
- Hermawati, Fajar Astuti. 2013. *Data Mining*. Yogyakarta: Andi.
- Sunjana. 2010. “Klasifikasi Data Nasabah Sebuah Asuransi Menggunakan Algoritma *C4.5*”. *Jurnal Aplikasi Teknologi Informasi 2010*.
- Anand, Sheila. Dkk. 2012. “Analysis of SEER Dataset for Breast Cancer Diagnosis using *C4.5* Classification Algorithm”. *International Journal of Advanced Research in Computer and*

Communication Engineering. 1, (2),
2278 – 1021.

K, Madhan, Dkk. 2012. “Classification Of Queries Using C4.5 Algorithm In Micro Blogging Supported Classroom”. *International Journal of Communications and Engineering*. 5, (5), 1-6.

Nazrian, Adli dan Paidi Hidayat. 2012. “Studi Tentang Keputusan Nasabah Dalam Menabung Di Bank Sumut Cabang Usu Medan Metode *Analytical Hierarchy Process (AHP)*”. *Jurnal studi tentang keputusan nasabah dalam menabung di bank*. 14-20.

Nailufarth, Quraratul A’yun. 2012. “Sistem Perbankan Dan Persoalan Riba Dalam Islam Menuju Sistem Perbankan Bebas Riba”. *Jurnal Sistem Perbankan Dan Persoalan Riba Dalam Islam Menuju Sistem Perbankan Bebas Riba*.

Sulistiyawaty, Risna. 2009. “Prilaku Konsumen Dalam Penggunaan Kartu Kredit Di Wilayah DKI Jakarta”. *Jurnal Prilaku Konsumen Dalam Penggunaan Kartu Kredit Di Wilayah DKI Jakarta*.