

## IMPLEMENTASI DATA MINING KLASIFIKASI C4.5 DALAM MENENTUKAN KELAYAKAN PENGAMBILAN KREDIT

Rika Nofitri<sup>1</sup>, Juna Eska<sup>2</sup>

<sup>1,2</sup>Sistem Informasi, STMIK Royal

email: <sup>1</sup>nofitririka15@gmail.com, <sup>2</sup>dosen.junaeska@gmail.com

**Abstrak:** *Data mining adalah serangkaian proses untuk menggali nilai tambah berupa pengetahuan dari suatu kumpulan data yang selama ini tidak diketahui secara manual. 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. Klasifikasi adalah membangun suatu model yang bisa mengklasifikasikan suatu objek berdasar atribut-atributnya. Klasifikasi adalah proses untuk menemukan model atau fungsi yang menjelaskan atau membedakan kelas data, dengan tujuan untuk dapat memperkirakan kelas dari suatu objek yang labelnya tidak diketahui. Pada penelitian ini akan dibahas teknik klasifikasi yang diterapkan untuk kelayakan pengambilan kredit. Data-data yang ada di Koperasi akan diolah untuk menghasilkan pohon keputusan yang berisi informasi tentang persyaratan yang harus dilengkapi oleh seorang calon nasabah dalam mengajukan kredit. Teknik klasifikasi yang akan digunakan adalah decision tree, Algoritma C4.5.*

**Kata kunci:** *Data mining, Decision tree, C4.5, Kredit, Koperasi*

### PENDAHULUAN

Beberapa tahun belakangan ini kemajuan dan perkembangan Bank Perkreditan Rakyat secara kuantitatif meningkat. Bank Perkreditan Rakyat atau lebih dikenal dengan istilah KOPERASI adalah salah satu jenis bank yang dikenal melayani golongan pengusaha mikro, kecil dan menengah dengan lokasi yang pada umumnya dekat dengan tempat masyarakat yang membutuhkan. Salah satu pelayanan yang diberikan KOPERASI adalah pemberian kredit kepada nasabah. Proses pemberian kredit dilakukan melalui berbagai pertimbangan dan persyaratan yang harus dipenuhi oleh nasabah, kemudian pihak KOPERASI mempertimbangkan apakah nasabah tersebut layak mengambil kredit atau tidak. Proses pemberian kredit di Koperasi Riama Kisaran, Kabupaten 50 Kota memerlukan waktu yang lama karena masih dilakukan secara manual, sehingga seringkali membuat nasabah berpindah ke perusahaan pembiayaan yang lain. Waktu yang terlalu lama tersebut disebabkan oleh proses analisa yang kurang akurat, sehingga harus dilakukan pengulangan dalam proses penilaiannya. Hal ini dirasa sangat tidak efisien dan efektif. Oleh sebab itu perlu dibuat suatu sistem pendukung kepu-

utusan kelayakan pengambilan kredit di Koperasi Riama Kisaran. Penggunaan *data mining* klasifikasi C4.5 akan menghasilkan suatu pohon keputusan, sehingga bila ada data masukan, dengan pohon keputusan akan didapatkan *output* yang jelas sebagai kebijakan pengambilan keputusan dalam pemberian kredit kepada nasabah.

Menurut jurnal Rika Nofitri 2017, algoritma C4.5 juga digunakan dalam pembentukan pohon keputusan untuk memprediksi pencapaian profit pada perusahaan distributor. Data yang digunakan adalah laporan laba rugi dari tahun 2012 hingga 2016 dengan jumlah 60 data. Dari penelitian tersebut dapat diketahui bahwa kriteria biaya menjadi kriteria yang paling penting dalam menentukan profit tercapai atau tidak.

Berdasarkan latar belakang masalah diatas, ujuan yang hendak dicapai dari penelitian ini adalah

1. Mendapatkan sebuah sistem pendukung keputusan kelayakan pengambilan kredit yang menggunakan *data mining* klasifikasi C4.5.
2. Mendapatkan sebuah pohon keputusan yang menggambarkan data dalam bentuk tabel yang kompleks menjadi bentuk pohon lebih *simple*.

3. Mendapatkan sebuah *rule* yang menjelaskan aturan kelayakan pengambilan kredit.

## METODOLOGI

Metodologi diawali dengan analisis masalah sampai pengujian terhadap analisa data nasabah dengan tools yang digunakan untuk menentukan kelayakan pengambilan kredit. Kerangka kerja yang dilakukan adalah sebagai berikut:

1. Identifikasi masalah

Permasalahan yang dihadapi Koperasi Riama Kisaran yaitu ketidakefisienan dalam memutuskan kelayakan pemberian kredit kepada nasabah. Sehingga menyebabkan nasabah berpindah pada lembaga peminjaman lain.

2. Analisis kebutuhan data

Ada dua jenis data yang dikumpulkan yaitu data analisa kredit dan rekapitulasi nama-nama nasabah yang pengajuan kreditnya diterima dan ditolak pada bulan Januari-Desember 2017.

Analisa kredit adalah formulir isian yang berisi syarat-syarat yang diisi oleh calon nasabah yang mengajukan pengambilan kredit dan kemudian akan dianalisis oleh bagian kredit untuk pengambilan keputusan apakah calon nasabah tersebut layak atau tidak diberikan kredit. Data selanjutnya adalah data rekapitulasi nama-nama nasabah yang pengajuan kreditnya diterima dan ditolak. Data ini akan diolah untuk pengujian metode klasifikasi yang dibangun.

3. Pengolahan data

Setelah dilakukan pengumpulan data, langkah selanjutnya adalah melakukan *filtering* terhadap persyaratan yang diperlukan untuk melakukan pengambilan kredit. Dari beberapa persyaratan yang terdapat didalam analisa kredit, diambil lima persyaratan penting yang akan dijadikan sebagai pembentuk pohon keputusan dalam metode klasifikasi. Adapun persyaratannya adalah syarat administrasi, jaminan, watak, surat izin usaha dan jenis pinjaman. Kriteria

persyaratan tersebut berpedoman pada peraturan yang telah ditetapkan di Koperasi Riama Kisaran. Kemudian dibentuk tabel yang berisi variabel input dan variabel target untuk diolah dengan metode klasifikasi.

4. Implementasi sistem

Setelah terbentuk tabel, langkah selanjutnya yaitu membentuk pohon keputusan dengan metode klasifikasi algoritma C4.5. Pembentukan pohon keputusan dimulai dengan memilih atribut sebagai akar. Atribut yang dijadikan akar adalah atribut yang memiliki nilai *gain* tertinggi dari perhitungan terhadap masing-masing atribut yang ada. Setelah didapatkan salah satu atribut yang dijadikan akar, langkah selanjutnya yaitu membuat cabang pada salah satu nilai akar. Akar yang dibentuk cabang adalah nilai perhitungan entropinya yang lebih dari nol. Begitu seterusnya sampai semua kasus memiliki kelas yang sama.

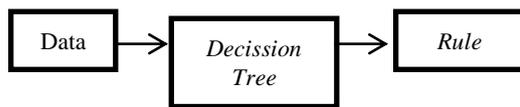
5. Pengujian Sistem

Pengujian sistem dilakukan menggunakan aplikasi rapid miner. Data diinputkan dalam bentuk tabel kemudian diproses sehingga membentuk pohon keputusan.

## Data Mining Klasifikasi C4.5

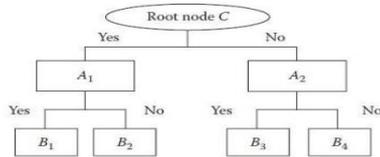
*Data Mining* adalah kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola atau hubungan dalam data berukuran besar. Keluaran *Data Mining* dapat digunakan untuk membantu pengambilan keputusan di masa depan.

Cara kerja teknik klasifikasi yaitu menyimpulkan karakteristik dari suatu kelompok tertentu. Metode ini meliputi penetapan suatu set data dengan satu set kelas yang telah diketahui, dan memetakan semua item lain ke dalam set data tersebut. Salah satu hasil dari metode klasifikasi yang sangat kuat dan terkenal yaitu pohon keputusan. Metode pohon keputusan mengubah fakta yang sangat besar menjadi pohon keputusan yang merepresentasikan aturan.



Gambar 1. Konsep Pohon Keputusan

Lebih lanjut bentuk pohon keputusan dapat dilihat pada gambar dibawah ini:



Gambar 2. Konsep Dasar Pohon Keputusan

Banyak algoritma yang dapat dipakai dalam pembentukan pohon keputusan antara lain ID3, CART, dan C4.5. Algoritma C4.5 merupakan pengembangan dari ID3. Oleh karena pengembangan tersebut algoritma C4.5 mempunyai prinsip dasar kerja yang sama dengan algoritma ID3. Hanya saja dalam algoritma C4.5 pemilihan atribut dilakukan dengan menggunakan *Gain Ratio*. Secara umum algoritma C4.5 untuk membangun pohon keputusan adalah sebagai berikut:

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 memilih atribut sebagai akar, didasarkan pada nilai *gain* tertinggi dari atribut-atribut yang ada. Untuk menghitung *gain* digunakan rumus sebagai berikut :

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

Sedangkan perhitungan nilai *entropy* dapat dilihat pada rumus berikut ini :

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

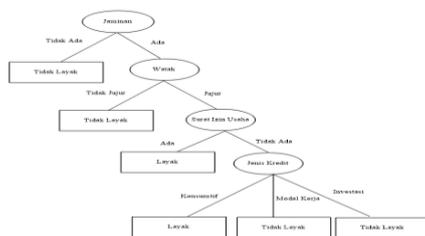
## HASIL DAN PEMBAHASAN

Berdasarkan data yang ada, digunakan sampel sebanyak dua puluh buah data untuk pengolahan secara manual. Data sampel yang akan diolah dapat dilihat pada tabel berikut.

Tabel 1. Sampel pengolahan data manual

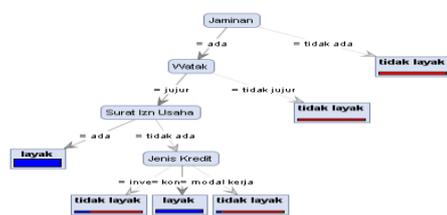
Nasabah	Syarat adm	Jaminan	Watak	Surat Izin Usaha	Jenis Kredit	Keputusan
1	Lengkap	Ada	Jujur	Tidak ada	Konsumtif	Layak
2	Lengkap	Ada	Jujur	Tidak ada	Modal Kerja	Tidak Layak
3	Lengkap	Ada	Jujur	Tidak ada	Investasi	Tidak Layak
4	Lengkap	Ada	Jujur	Ada	Konsumtif	Layak
5	Belum Lengkap	Ada	Jujur	Ada	Investasi	Layak
6	Belum Lengkap	Ada	Jujur	Ada	Modal Kerja	Layak
7	Lengkap	Tidak Ada	Jujur	Ada	Modal Kerja	Tidak Layak
8	Lengkap	Tidak Ada	Jujur	Ada	Konsumtif	Tidak Layak
9	Lengkap	Tidak Ada	Jujur	Ada	Investasi	Tidak Layak
10	Lengkap	Ada	Tidak Jujur	Ada	Investasi	Tidak Layak

11	Belum Lengkap	Ada	Jujur	Ada	Modal Kerja	Layak
12	Lengkap	Tidak Ada	Jujur	Ada	Konsumtif	Tidak Layak
13	Lengkap	Ada	Tidak Jujur	Ada	Modal Kerja	Tidak Layak
14	Belum Lengkap	Ada	Tidak Jujur	Ada	Investasi	Tidak Layak
15	Lengkap	Ada	Jujur	Ada	Investasi	Layak
16	Belum Lengkap	Ada	Jujur	Tidak ada	Modal Kerja	Tidak Layak
17	Belum Lengkap	Ada	Jujur	Tidak ada	Konsumtif	Layak
18	Lengkap	Ada	Jujur	Ada	Modal Kerja	Layak
19	Lengkap	Tidak Ada	Tidak Jujur	Ada	Investasi	Tidak Layak
20	Lengkap	Ada	Jujur	Ada	Modal Kerja	Layak



Gambar 3. Hasil Pohon Keputusan

Hasil klasifikasi yang telah dijalankan pada *software Rapidminer* dapat dilihat pada gambar dibawah ini:



Gambar 4. Hasil Keputusan

Berdasarkan pohon keputusan diatas dapat diketahui ada 4 kriteria yang mempengaruhi kelayakan pengambilan kredit. Kriteria jaminan menjadi *node* akar. Sedangkan atribut syarat adm tidak menjadi

kriteria yang mempengaruhi pengambilan kredit.

## SIMPULAN

Berdasarkan hasil penelitian dan pembahasan yang telah dilakukan, maka dapat ditarik kesimpulan yaitu bentuk pohon keputusan dalam memberikan rekomendasi terhadap pengambilan keputusan untuk kelayakan pengambilan kredit adalah dilihat kriteria pertama yang menjadi *root* yaitu jaminan. Kriteria jaminan mempunyai *sub root* kedua yaitu watak. Kemudian *subroot* dari watak yaitu surat izin usaha. Selanjutnya *sub root* yang terakhir yaitu jenis kredit. Rule yang dihasilkan dapat dijadikan sebagai acuan dalam memutuskan kelayakan pemberian kredit kepada calon nasabah sehingga pengambilan keputusan dapat dilakukan dengan efektif dan efisien.

## DAFTAR PUSTAKA

- Sembiring, M. A. (2016). *Penerapan metode decision tree algoritma c45 untuk memprediksi hasil belajar mahasiswa berdasarkan riwayat akademik. JURTEKSI ROYAL Vol 3 No 1, 3.*
- Eska, J. (2018). Penerapan Data Mining Untuk Prediksi Penjualan Wallpaper Menggunakan Algoritma C4. 5.
- Nofitri, Rika. "Analisa Kinerja Algoritma C. 45 Dalam Memprediksi Pencapaian Profit." (2017)
- Arfyanti, I., & Purwanto, E. (2012). *Aplikasi Sistem Pendukung Keputusan Pemberian Kelayakan Kredit Pinjaman Pada Bank Rakyat Indonesia Unit Segiri Samarinda Dengan Metode Fuzzy MADM (Multiple Attribute Decision Making) Menggunakan SAW (Simple Additive Weighting). Semantik, 2(1).*