

Penerapan Algoritma C4.5 Dalam Menentukan Pemberian Pinjaman Bagi Nasabah Di CU.Mitra Kasih Smart

Hanson Apriando Sibarani¹, Indra M Sarkis², Harlen Gilbert Simanullang³
^{1,2,3}Fakultas Ilmu Komputer, Universitas Methodist Indonesia

Info Artikel

Histori Artikel:

Received, Des 12, 2022
Revised, Jan 20, 2023
Accepted, Apr 21, 2023

Keywords:

Pimberian, Pinjaman, Nasabah, C4.5

ABSTRAK

Koperasi simpan pinjam CU. Mitra Kasih Smart adalah badan usaha yang bergerak dibidang simpan pinjam salah satu usaha pada koperasi itu adalah pemberian kredit. Masalah yang dihadapi koperasi simpan pinjam saat ini dalam pemberian kredit adalah resiko terlambatnya pembayaran, pengembalian bahkan kegagalan pembayaran kredit. Masalah ini terjadi karena penyalahgunaan kredit dan lemahnya pengawasan baik dalam proses pemberian kredit maupun dalam palaksanaannya. Solusi yang tepat untuk memecahkan permasalahan yang ada yaitu dengan menggunakan algoritma data mining. Konsep data mining akan mempermudah mengatasi masalah yang belum optimal di koperasi, metode klasifikasi mampu menemukan model yang membedakan konsep atau kelas data dengan tujuan untuk mempermudah dalam memprediksi peluang dimasa depan berdasarkan pengalaman dimasa sebelumnya. Penulis melakukan penelitian ini pada koperasi CU Mitra Kasih Smart Bagan Batu dengan “ Penerapan Algoritma C4.5 dalam menentukan pemberian pinjaman bagi nasabah di CU. Mitra Kasih Smart”. Dalam penelitian ini penulis menggunakan 7 atribut sebagai penilaian yaitu: Nama, jumlah simpanan, penghasilan, jumlah pinjaman, lama pinjaman. Pengujian dilakukan dengan menggunakan 162 data. Dari hasil pengujian yang dilakukan dengan menggunakan *tools RapidMiner* dapat disimpulkan bahwa tingkat *accuracy* sebesar 77,78%.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Penulis Koresponden:

Hanson Apriando Sibarani,
Fakultas Ilmu Komputer,
Universitas Methodist Indonesia, Medan,
Jl. Hang Tuah No.8, Medan - Sumatera Utara.
Email: ashanson08@gmail.com

1. PENDAHULUAN

Koperasi adalah badan usaha beranggotakan orang perseorangan atau badan hukum koperasi. Kegiatan koperasi didasarkan pada prinsip-prinsip koperasi, dan gerakan sosial ekonomi berdasarkan kekerabatan. Melalui gerakan koperasi, peningkatan kemampuan kerjasama akan sangat menekan angka pengangguran. Perihal perkoperasian diatur dengan Undang-Undang Nomor 25 Tahun 1992. Adanya perubahan dari Undang-Undang Nomor 25 Tahun 1992 menjadi UU 17 tahun 2012 mengenai perkoperasian, mencakup reformasi hukum agar koperasi dapat terwujud untuk organisasi ekonomi yang sehat, kuat, mandiri dan tangguh beserta badan usaha yang andal dalam kegiatan didasarkan pada nilai dan prinsip koperasi. Peran koperasi dapat membuktikan eksistensinya dalam menghadapi krisis sosial dan ekonomi saat ini. Ini merupakan bagian yang dapat mendorong bagi yang berminat jadi keanggotaan koperasi.

Koperasi CU Mitra Kasih Smart memiliki fasilitas seperti pemberian pinjaman atau kredit. Nasabah yang mengajukan pinjaman atau kredit memiliki kondisi ekonomi yang berbeda-beda, sehingga nasabah berpotensi tidak dapat melunasi angsuran. Hal tersebut terjadi dikarenakan banyaknya nasabah yang mengajukan pinjaman, penilaian tidak objektif karena ada simpati dan empati oleh tim penilai dan keterbatasan sumber daya manusia yang mengelolah CU untuk menghindari kehilapan atau kesilapan dalam menentukan layal atau tidak layakannya nasabah meminjam, hal tersebut lah yang memnyebabkan prosedur pemberian pinjaman menyalahi aturan sehingga terjadi kredit macet

Sampai saat ini pihak koperasi masih menggunakan keputusan personal yang berdasarkan jaminan saja dan belum memiliki sistem dalam menentukan apakah peminjam itu layak atau tidak diberikan pinjaman sehingga banyak terjadi penyimpangan standart operasional perosedur pinjaman nasabah. Hal tersebut dikarenakan keterbatasan dalam melibatkan analis kredit yang membutuhkan biaya cukup mahal dan dapat memberatkan nasabah, kesederhanaan inilah yang memicu terjadinya kredit macet dan dapat menyebabkan kerugian pada koperasi CU Mitra Kasih Smart.

Pada penelitian ini penulis menerapkan metode algoritma data mining seperti Algoritma C4.5 untuk melakukan pengambilan keputusan pemberian kredit pada nasabah di koperasi CU Mitra Kasih Smart. Dengan diterapkannya metode Algoritma C4.5 maka diharapkan koperasi CU Mitra Kasih Smart dapat membuat suatu klasifikasi pemberian kredit kepada nasabah berdasarkan riwayat data pinjaman sebelumnya serta dapat memanfaatkan data karakteristik nasabah untuk dijadikan sebagai informasi tambahan yang dapat membantu koperasi dalam membuat keputusan dalam memberikan pinjaman.

2. METODE PENELITIAN

2.1 Decision Tree

Merupakan metode klasifikasi yang paling sering digunakan. Karena bentuk strukturnya yang hirarki atau menggunakan struktur pohon dengan proses mengubah bentuk data menjadi model pohon, mengubah bentuk pohon menjadi rule dan menyederhanakan rule. [1], [2]

2.2 Algoritma C4.5

Algoritma C4.5 merupakan salah satu teknik Decision tree yang sering digunakan, yang menghasilkan beberapa aturan-aturan dan sebuah pohon keputusan dengan tujuan untuk meningkatkan keakurtan dari prediksi yang sedang dilakukan, di samping itu algoritma C4.5 merupakan algoritma yang mudah dimengerti.[3], [4]

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

1. Pilih atribut sebagai akar.
2. Buat cabang untuk tiap-tiap nilai
3. Bagi kasus dalam cabang
4. Ulangi proses untuk setiap cabang

Tahap awal ialah penetapan akar dalam pohon keputusan. Pada pemilihan atribut untuk akar di algoritma C4.5, didasari oleh skor gain paling tinggi daripada atribut-atributnya pada dataset. Dalam penghitungan skor gain bisa dilihat dirumus 1 seperti dibawah ini:

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

Keterangan: S = Himpunan kasus |S_i| = Jumlah kasus pada partisi ke-I |S| = Jumlah kasus dalam S
A = Atribut N = Jumlah partisi atribut A

Saat menghitung skor gain, tetapkan terlebih dahulu skor entropy pada setiap atribut yang ada. Gain mempunyai skor paling tinggi sehingga menjadi akar. Lakukan Langkah yang sama untuk mencari skor gain berikutnya dengan mencari node selanjutnya. Penghitungan skor entropy pada rumus 2 seperti dibawah ini:

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

Keterangan: S= Himpunan kasus , A = Atribut, n = Jumlah partisi S, pi = Proporsi dari Si terhadap S

3. HASIL DAN PEMBAHASAN

Pada penelitian ini data yang digunakan adalah data sekunder atau data yang berasal dari jurnal, paper, buku serta beberapa informasi lainnya yang berhubungan dengan penelitian yang dikutip. Pengolahan data yang dibutuhkan dimulai dengan menentukan pengelompokan data dengan menentukan himpunan dan atribut, berdasarkan data pada CU.Mitra Kasih Smart

Tabel 3. 1 Klasifikasi Atribut Penghasilan

PENGHASILAN	
<5 Juta	Sedikit
5 – 10 Juta	Cukup
>10 Juta	Banyak

Tabel 3. 2 Klasifikasi Atribut Jumlah Pinjaman

JUMLAH PINJAMAN	
<10 Juta	Rendah
10 – 25 Juta	Sedang
>25 Juta	Tinggi

Tabel 3.3 Klasifikasi Atribut Lama Pinjaman

LAMA PINJAMAN	
<2 Tahun	Pendek
2 – 4 Tahun	Cukup Panjang
4 < 5 Tahun	Panjang

Berikut adalah data yang digunakan :

Tabel 3. 4 Data yang Belum Diklasifikasi

No	Nama	Jumlah Simpanan	Penghasilan	Agunan	Jumlah Pinjaman	Lama Pinjaman	Status
1	Putra Sinaga	Rp2.000,000	Rp4.500,000	Rp5.000,000	Rp3.500,000	3	Layak
2	Poniatik	Rp10.000,000	Rp2.500,000	Rp3.000,000	Rp3.000,000	10	Layak
3	Geradus Bayu	Rp2.300,000	Rp3.000,000	Rp5.000,000	Rp8.000,000	6	Tidak Layak
4	Sri Lestari	Rp2.100,000	Rp4.000,000	Rp4.400,000	Rp7.000,000	6	Tidak Layak
5	Rido Prasetya	Rp7.000,000	Rp4.000,000	Rp5.000,000	Rp10.000,000	12	Layak
6	Soter manurung	Rp20.000,000	Rp7.000,000	Rp5.000,000	Rp8.000,000	6	Layak
.
.
.
162	Marjoko sihombing	Rp2.400,000	Rp3.000,000	Rp5.000,000	Rp11.000,000	12	Tidak Layak

3.1 Penguian Model

1. Entropy

Menghitung nilai total kasus keseluruhan yang ada berdasarkan jumlah kasus yang layak dan tidak layak. Kemudian akan diproses dengan mencari nilai Entropy dari setiap atribut dapat ditunjukkan pada rumus[5]

a. Jumlah Simpanan

$$\text{Entropy (SS)} = (-\frac{44}{80}) \cdot \log_2(\frac{44}{80}) + (-\frac{36}{80}) \cdot \log_2(\frac{36}{80})$$

$$= 0.992774454$$

$$\text{Entropy (SSE)} = (-\frac{40}{50}) \cdot \log_2(\frac{40}{50}) + (-\frac{10}{50}) \cdot \log_2(\frac{10}{50})$$

$$= 0.721928095$$

$$\text{Entropy (SB)} = (-\frac{26}{32}) \cdot \log_2(\frac{26}{32}) + (-\frac{6}{32}) \cdot \log_2(\frac{6}{32})$$

$$= 0.757878463$$

b. Penghasilan

$$\begin{aligned} \text{Entropy (S)} &= (-(58/92)*\log_2(58/92))+(-(34/92)*\log_2(34/92)) \\ &= 0.958348517 \end{aligned}$$

$$\begin{aligned} \text{Entropy (C)} &= (-(37/54)*\log_2(37/54))+(-(17/54)*\log_2(17/54)) \\ &= 0.898653376 \end{aligned}$$

$$\begin{aligned} \text{Entropy (B)} &= (-(15/16)*\log_2(15/16))+(-(1/16)*\log_2(1/16)) \\ &= 0.337290067 \end{aligned}$$

c. Agunan

$$\begin{aligned} \text{Entropy (K)} &= (-(80/112)*\log_2(80/112))+(-(32/112)*\log_2(32/112)) \\ &= 0.863120569 \end{aligned}$$

$$\begin{aligned} \text{Entropy (B)} &= (-(30/50)*\log_2(30/50))+(-(20/50)*\log_2(20/50)) \\ &= 0.981453895 \end{aligned}$$

d. Jumlah Pinjaman

$$\begin{aligned} \text{Entropy (PE)} &= (-(90/134)*\log_2(90/134))+(-(44/134)*\log_2(44/134)) \\ &= 0.925685987 \end{aligned}$$

$$\begin{aligned} \text{Entropy (LP)} &= (-(13/21)*\log_2(13/21))+(-(8/21)*\log_2(8/21)) \\ &= 0.958711883 \end{aligned}$$

$$\begin{aligned} \text{Entropy (PA)} &= (-(9/9)*\log_2(9/9))+(-(0/9)*\log_2(0/9)) \\ &= 0 \end{aligned}$$

2. Information Gain

Perhitungan nilai Gain pada setiap himpunan atribut, dibawah ini proses perhitungannya. Menghitung nilai Gain dapat ditunjukkan pada rumus di atas.[5]

a. Gain (Total, Jumlah Simpanan)

$$\text{Gain}(S, A) = \text{Entropy}(S)$$

$$- \sum_{i=1}^n \frac{|\text{Status tempat tinggal}|}{|\text{Total}|} * \text{Entropy}(\text{status tempat tinggal})$$

$$0.91199892 - \left(\frac{80}{162} * 0.992774454 + \frac{50}{162} * 0.721928095 + \frac{32}{162} * 0.757878463\right) = 0.049218229$$

b. Gain (Total, Penghasilan)

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

$$0.91199892 - \left(\frac{92}{162} * 0.958348517 + \frac{54}{162} * 0.898653376 + \frac{16}{162} * 0.337290067\right) = 0.034887889$$

c. Gain (Total, Agunan)

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

$$0.91199892 - \frac{112}{162} * 0.863120569 + \frac{50}{162} * 0.981453895 = 0.012355719$$

d. Gain (Total, Jumlah Pinjaman)

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

$$0.91199892 - \frac{89}{162} * 0.840980534 + \frac{45}{162} * 0.894451885 + \frac{28}{162} * 0.99631652 = 0.029317038$$

e. Gain (Total, Lama Pinjaman)

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

$$0.91199892 - \frac{134}{162} * 0.925685987 + \frac{21}{162} * 0.958711883 + \frac{9}{162} * 0 = 0.033458797$$

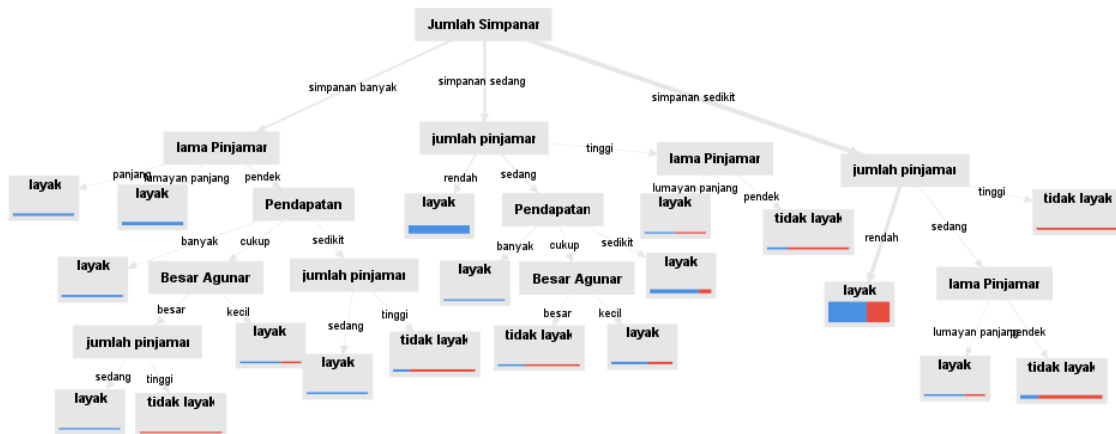
Berikut adalah tabel nilai Entropy dan Gain:

Tabel 3. 5 Nilai Entropi dan Gain

No	Atribut	Value	Jumlah Kasus	Layak	Tidak Layak	Entropy	Gain
1	Total		162	109	53	0.91199892	
2	Jumlah simpanan	Simpanan sedikit	80	44	36	0.992774454	0.049218229
		Simpanan sedang	50	40	10	0.721928095	
		Simpanan banyak	32	26	6	0.757878463	
3	Penghasilan	Sedikit	92	58	34	0.958348517	0.034887889
		Cukup	54	37	17	0.898653376	
		Banyak	16	15	1	0.337290067	
4	Agunan	Kecil	113	80	32	0.863120569	0.012355719
		Besar	50	30	20	0.981453895	
5	Jumlah Pinjaman	Rendah	89	65	24	0.840980534	0.029317038
		Sedang	45	31	14	0.894451885	
		Tinggal	28	14	14	0.99631652	
6	Lama Pinjaman	Pendek	134	90	44	0.925685987	0.033458797
		Lumayan Panjang	21	13	8	0.958711883	
		Panjang	9	9	0	0	

3.1. Bentuk Pohon Keputusan

Dari Tabel 3.5 tersebut dapat digambarkan pohon keputusan seperti gambar 3. 1



Gambar 3.1 Pohon Keputusan

Berdasarkan data set, klasifikasi penentuan kelayakan ini dilakukan berdasarkan data klasifikasi anggota koperasi karyawan yang telah tersedia. Pohon keputusan algoritma C4.5 dibentuk melalui penentuan nilai gain dengan perolehan tertinggi dari setiap menghitung node, didasarkan node yang ditentukan, maka untuk klasifikasinya dapat dilihat pada gambar 3.1.

Perolehan hasil pengujian model sudah dilaksanakan menggunakan algoritma C4.5 dilaksanakan pengujian dengan level akurasi mempergunakan confusion matrix menggunakan aplikasi Rapidminer dengan cara memasukkan data yang didapatkan dari CU.Mitra Kasih Smart. Untuk perolehan hasil evaluasi sebagai berikut:

3.2. Confusion Matrix

Confusion Matrix adalah pengukuran performa untuk masalah klasifikasi machine learning dimana keluaran dapat berupa dua kelas atau lebih. Confusion Matrix adalah tabel dengan 4 kombinasi berbeda dari nilai prediksi dan nilai aktual. Ada empat istilah yang merupakan representasi hasil proses klasifikasi pada confusion matrix yaitu True Positif, True Negatif, False Positif, dan False Negatif. [6]

Berdasarkan data perusahaan didapatkan dari 162 pengajuan pinjaman terdapat 109 pinjaman tersebut layak dan 53 tidak layak tetapi berdasarkan prediksi Algoritma C4.5 terdapat 135 layak dan 27 tidak layak .

Tabel 3.1 Confusion Matrix

N= 162	Aktual layak	Aktual tidak layak
Prediksi Layak	104	31
Predik tidak layak	5	22
	109	53

Penghitungan akurasi menggunakan data training dengan algoritma C4.5. Dengan mengetahui data training sebanyak 162 record data, Hasil accuracy yang di peroleh dari data training mencapai 77,78%. Dari data tersebut diketahui prediksi Layak dengan true layak mencapai 104 dan true tidak layak mencapai 22, sedangkan untuk false true mencapai 31 dan false tidak layak mencapai 5.

Perolehan hasil untuk confusion matrix ditabel 3.1, untuk langkah berikutnya dilaksanakan penghitungan untuk skor accuracy, precision, dan recall. Perbandingan skor accuracy, precision, dan recall yang sudah dilakukan penghitungan pada metode algoritma C4.5.

4. KESIMPULAN

Berdasarkan hasil pengujian dan penelitian penerapan metode algoritma C4.5 dalam menentukan pemberian pinjaman di CU.Mitra kasih Smart yang telah dilakukan, maka dapat disimpulkan sebagai berikut :

Analisa penggunaan data mining dengan algoritma C4.5 terbukti akurat dalam penentuan pemberian kredit. Hal ini dibuktikan dengan hasil evaluasi penelitian bahwa algoritma C4.5 mendapatkan nilai akurasi sebesar 77.78%

REFERENSI

- [1] Ainurrohmah, "Akurasi Algoritma Klasifikasi pada Software Rapidminer dan Weka," *Prisma*, vol. 4, pp. 493–499, 2021.
- [2] K. Kelvin, B. Mulyawan, and T. Sutrisno, "Penerapan Algoritma C4.5 Untuk Penerimaan Kelayakan Kredit Pada Koperasi (Studi Kasus:Koperasi Xyz)," *J. Ilmu Komput. dan Sist. Inf.*, vol. 8, no. 1, p. 109, 2020, doi: 10.24912/jiksi.v8i1.11478.
- [3] Y. Kurnia and E. D. Kusuma, "Pendugaan Kredit Macet Pada Koperasi Simpan Pinjam Flamboyan Binaan PPSW Jakarta Dengan Menggunakan Komparasi Algoritma Naïve Bayes dan C4. 5," *Algor*, vol. 1, pp. 66–74, 2020.
- [4] E. Sugiyarti, K. A. Jasmi, B. Basiron, M. Huda, K. Shankar, and A. Maseleno, "Decision support system of scholarship grantee selection using data mining," *Int. J. Pure Appl. Math.*, vol. 119, no. 15, 2018, doi: 10.5772/47788.
- [5] S. D. Manullang, E. Buulolo, and I. Lubis, "Implementasi Data Mining Dalam Memprediksi Jumlah Pinjaman Dengan Algoritma C4.5 Pada Kopdit CU Damai Sejahtera," *J. Sist. Komput. dan Inform.*, vol. 1, no. 3, p. 265, 2020, doi: 10.30865/json.v1i3.2153.
- [6] I. Menarianti, "Klasifikasi data mining dalam menentukan pemberian kredit bagi nasabah koperasi," *J. Ilm. Teknosains*, vol. 1, no. 1, pp. 1–10, 2015.