

Penentuan Kelayakan Kredit Menggunakan Algoritma C4.5 Pada Koperasi Karyawan PT. HPPM Karawang

Siti Rahma Alfadillah

Universitas Bina Sarana Informatika

Jl. Kramat Raya No. 98, RT.2RW.9, Kwitang, Kec. Senen, Kota Jakarta Pusat, Daerah Khusus
Ibukota Jakarta 10450, Indonesia

Email: 19236027@bsi.ac.id

Submit: 00-00-0000 | Revisi : 00-00-0000 | Terima : 00-00-0000 | Publikasi: 00-00-0000

Abstrak

Koperasi karyawan merupakan sebuah instansi yang berada dibawah perusahaan yang semua anggota dan pengurus nya adalah karyawan di perusahaan itu sendiri. Salah satu fasilitas yang disediakan oleh koperasi ini adalah jasa pinjaman, yang selalu dibutuhkan oleh para anggota nya, karena anggota koperasi dapat meminjam sejumlah uang dengan cicilan dan margin yang ringan. Pada Koperasi PT. HPPM ini cicilan perbulannya tidak boleh melebihi 40% dari gaji pokok, akan tetapi kenyataannya masih banyak anggota yang cicilan perbulannya melebihi 40% dari gaji pokok. Oleh karena itu pada penelitian ini bertujuan untuk mengklasifikasikan kelayakan anggota dalam kategori layak dan tidak layak dengan cara melakukan analisis persyaratan yang ditetapkan atas pengajuan kredit menggunakan metode algoritma c4.5 yang menghasilkan pohon keputusan dan pengujian yang dilakukan dengan aplikasi Rapidminer. Dalam penelitian ini menggunakan 7 atribut, diantaranya masa anggota, jenis pinjaman, nominal pengajuan, jangka waktu, gaji pokok, payroll, over limit dan 1 class label. Hasil dari penelitian ini didapatkan akurasi sebesar 97.02% artinya akurasi menggunakan algoritma c4.5 terbilang sangat tinggi. Diharapkan hasil dari penelitian ini dapat dijadikan sebagai acuan untuk pihak koperasi dalam menentukan anggota yang layak mendapatkan kredit.

Kata Kunci: Algoritma C4.5, Klasifikasi, Koperasi, Kredit

Abstract

An employee cooperative is an agency under a company whose members and administrators are all employees of the company itself. One of the facilities provided by this cooperative is loan services, which its members always need, because cooperative members can borrow a certain amount of money with low installments and margins. At Koperasi PT. HPPM's monthly installments cannot exceed 40% of the basic salary, but in reality there are still many members whose monthly installments exceed 40% of the basic salary. Therefore, this research aims to classify the eligibility of members into eligible and non-eligible categories by analyzing the requirements set for credit applications using the c4.5 algorithm method which produces a decision tree and testing is carried out using the Rapidminer application. In this research, 7 attributes are used, including membership period, type of loan, nominal application, term, basic salary, payroll, over limit and 1 class label. The results of this research obtained an accuracy of 97.02%, meaning that the accuracy using the c4.5 algorithm is said to be very high. It is hoped that the results of this research can be used as a reference for cooperatives in determining members who deserve credit.

Keywords: C4.5 Algorithm, Classification, Cooperatives, Credit

1. Pendahuluan

Koperasi merupakan salah satu lembaga keuangan yang melakukan transaksi kredit. Dalam rangka mengembangkan usahanya dalam bentuk pinjaman anggota, koperasi memiliki prinsip kehati-hatian sebagaimana yang diterapkan pada perbankan dengan melakukan analisis pemberian kredit kepada calon anggotanya. (Lestari et al., 2020)

Koperasi karyawan merupakan sebuah instansi yang berada dibawah perusahaan yang semua anggota dan pengurus nya adalah karyawan di perusahaan itu sendiri. Salah satu fasilitas yang disediakan oleh koperasi ini adalah jasa pinjaman, yang selalu dibutuhkan oleh para anggota nya, karena anggota koperasi dapat meminjam sejumlah uang dengan cicilan dan margin yang ringan. (Riswanto & Laluma, 2020)

Melalui koperasi simpan pinjam anggota dapat melakukan pengajuan pinjaman dana ke koperasi dengan ketentuan yang sudah ditetapkan dan melakukan pembayaran angsuran dan bunga sesuai dengan yang ditentukan.



Koperasi dapat melayani beberapa transaksi diantaranya simpanan tabungan, tarik tabungan, dan pinjaman. Dalam transaksi pinjaman, pengelola koperasi harus melakukan penentuan kelayakan pinjaman, dimana pengelola harus menganalisa kelayakan kreditnya secara personal dengan mengisi formulir pengajuan pinjaman disertai dengan persyaratan tertentu. Meningkatnya pengajuan pinjaman, muncul dampak negatif bagi koperasi itu sendiri, dimana semakin lama pengelola melakukan analisa dalam hari-harinya, ditambah lagi dengan penentuan kelayakan pinjaman yang harus cair maksimal 24 jam dari waktu pengajuan, muncul kendala yang dirasakan koperasi dalam melakukan analisa kelayakan kredit tersebut, dimana ketika analisa pemberian kredit kurang tepat, akan menyebabkan over limit pinjaman. (Raihan M. R. et al., 2022) Oleh karena itu perlu evaluasi dalam analisa penentuan kelayakan kredit pada koperasi.

Penelitian ini berkaitan dengan koperasi simpan pinjam dalam menentukan pemberian kredit kepada para anggota koperasi. Koperasi simpan pinjam berperan memberikan pinjaman kepada para anggotanya dengan bunga yang telah ditentukan, akan tetapi kenyataannya masih terdapat adanya permasalahan, banyak anggota yang melebihi limit pinjaman 40% dari upah.

Permasalahan yang dihadapi tersebut haruslah segera diselesaikan dengan baik, jika permasalahan terus berlanjut dan semakin banyaknya pemberian kredit yang over limit akan menyulitkan koperasi untuk mengelola informasi keuangan dan merugikan bagi perusahaan.

Kopkar PT HPPM memiliki 1657 anggota dan yang aktif melakukan pinjaman sekitar 670 anggota. Untuk membantu pengelola dalam mengantisipasi adanya anggota yang melebihi limit pinjaman di Kopkar PT HPPM, pada penelitian ini akan dilakukan prediksi dengan melakukan simulasi terkait kelayakan anggota koperasi yang mendapatkan pinjaman berdasarkan sampel data yang ada di Kopkar PT. HPPM menggunakan Algoritma C4.5. Algoritma C4.5 merupakan algoritma yang umum digunakan untuk keperluan klasifikasi data dengan metode klasifikasi menggunakan pohon keputusan. Algoritma C4.5 berperan dalam pembentukan pohon keputusan sebagai akhir dalam menentukan prediksi kelayakan penerimaan kredit.

Metode algoritma C4.5 memiliki akurasi yang sangat baik, dibuktikan dengan penelitian yang berjudul Penentuan Kelayakan Penerimaan Pembiayaan Kredit Menggunakan Metode Naïve Bayes dan Algoritma C4.5 dimana pada penelitian ini akurasi yang di hasilkan dari metode algoritma c4.5 lebih baik dari metode Naïve bayes yaitu 91.23% dan AUC 0.686, maka dalam penelitian ini akan menerapkan algoritma C4.5 pada penentuan kelayakan penerimaan kredit pada Kopkar PT HPPM, diharapkan agar dapat meningkatkan kinerja Algoritma C4.5 dan menghasilkan keputusan yang akurat.

2. Metode

2.1. Tahapan Penelitian

Berikut ini adalah langkah-langkah yang digunakan dalam penelitian :



Sumber : (Sodikin et al., 2021)

Gambar 1. Tahapan Penelitian

1. Mengidentifikasi Masalah
Pada tahap awal penelitian ini adalah mengidentifikasi masalah apa yang terjadi pada Kopkar PT. HPPM. Masih adanya anggota yang kredit nya melebihi limit 40% dari upah menyulitkan koperasi untuk mengelola informasi keuangan.
2. Mempelajari Literatur
Setelah mengerti dan memahami permasalahan apa yang sedang terjadi, Langkah selanjutnya yaitu mencari referensi dan teori yang relevan dengan penelitian yang dilakukan. Pada penelitian ini menggunakan karya ilmiah dari beberapa jurnal dan buku sebagai bahan referensi.
3. Mengumpulkan Data
Proses mengumpulkan data dilakukan melalui observasi dan wawancara kepada karyawan Kopkar PT. HPPM untuk mengetahui data yang dibutuhkan dalam penelitian ini. Hasil yang diperoleh berupa data kredit anggota koperasi pada Januari – Mei 2024, maka nantinya data tersebut akan menjadi bahan pengujian dalam penelitian ini.

4. Menerapkan Metode
Berdasarkan hasil literatur dan mengenai permasalahan yang ada yaitu tentang anggota yang kreditnya sudah melebihi limit 40% dari upah pada koperasi. Algoritma C4.5 adalah metode yang dipilih untuk mengukur dan juga menganalisa seberapa akurat kelayakan pemberian kredit pada anggota di Kopkar PT.HPPM.
5. Mengelola Data
Pengolahan data ini menggunakan aplikasi Rapidminer melalui tahap pe-processing data untuk mengetahui apakah data terjadi missing value atau tidak.
6. Analisa Hasil
Pada tahap ini, dilakukan analisis hasil pengolahan data menggunakan perhitungan manual yang kemudian dicocokkan juga dengan pengujian menggunakan Aplikasi Rapid Miner. Fungsi penggunaan rapid miner ini guna mengukur dan mengetahui seberapa besar nilai akurasi data untuk menentukan kelayakan pemberian kredit kepada anggota di Kopkar PT. HPPM.

2.2. Decision Tree

“Decision tree atau seringkali diterjemahkan sebagai pohon Keputusan adalah salah satu algoritma machine learning yang bekerja mirip dengan cara manusia mengambil Keputusan. Algoritma decision tree akan membentuk flow chart berbentuk pohon terbalik, dimana akarnya (root) berada di atas sedangkan cabang dan daun (leaf) berada di bawah.” (Rifkie Primartha & Romi Satria Wahono, 2021)

2.2. Data Mining

Data Mining adalah suatu proses ekstraksi pengetahuan atau informasi yang berharga dari suatu set data yang besar dan kompleks, Tujuan utama dari data mining adalah mengidentifikasi pola, hubungan, atau pengetahuan yang berharga dan tersembunyi dalam suatu set data besar atau kompleks. Proses data mining bertujuan untuk menggali wawasan yang tidak dapat ditemukan secara langsung melalui pengamatan sederhana terhadap data. Manfaat utama dari data mining adalah pengambilan keputusan yang lebih baik, penemuan pola dan tren, prediksi masa depan, segmentasi pelanggan, efisiensi operasional, pengelolaan risiko, pemasaran yang lebih efektif dan peningkatan layanan pelanggan. (Prastyadi Wibawa Rahayu et al., 2024)

2.3. Klasifikasi

“Teknik classification bekerja dengan mengelompokkan data berdasarkan data training dan nilai atribut klasifikasi. Aturan pengelompokan tersebut akan digunakan untuk klasifikasi data baru ke dalam kelompok yang ada. Classification dapat direpresentasikan dalam bentuk pohon Keputusan (decision tree). Setiap node dalam pohon Keputusan menyatakan suatu tes terhadap atribut dataset, sedangkan setiap cabang menyatakan hasil dari tes tersebut” (Amril Mutoi Siregar & Adam Puspabhuana, 2017)

2.4. Algoritma C4.5

“Algoritma yang dipublikasikan pada tahun 1993 ini dapat digunakan untuk data bernilai kategorial maupun numerik (kontinu). Untuk data bernilai numerik, C4.5 membangun nilai-nilai batas (thresholds) lalu memilah data ke dalam sejumlah interval sehingga didapatkan nilai-nilai kategorial. Selain itu, C4.5 juga dapat menangani data dengan nilai-nilai atribut yang hilang (kosong)”. (Suyanto, 2018)

“Rumus 1 adalah rumus untuk menghitung nilai Entropy” (Najib et al., n.d.)

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

Keterangan :

S = Himpunan Kasus

n = Jumlah Partisi dalam atribut

pi = Proporsi dari Si terhadap S

Untuk menghitung nilai Gain dapat menggunakan rumus 2

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

Keterangan:

S = himpunan kasus

A = atribut

n = jumlah partisi atribut A

|Si| = jumlah kasus pada partisi ke – i

|S| = jumlah kasus dalam S

3. Hasil dan Pembahasan

3.1. Pengumpulan Data

Dalam melakukan penelitian, data merupakan sebuah objek atau pendukung untuk digunakan dalam penelitian. Sumber data pada penelitian ini adalah data anggota yang melakukan kredit pada Kopkar PT. HPPM dari bulan Januari – Mei 2024. Data yang di dapat ini melalui observasi pada Kopkar PT. HPPM dan melakukan wawancara pada karyawan Kopkar PT. HPPM.

Setelah mendapatkan data, Langkah pertama yang dilakukan adalah mengakumulasi data, dalam penelitian ini menggunakan data sekunder, yaitu data yang di dapat langsung dari Kopkar PT. HPPM. Kemudian data yang di dapat akan di jadikan dataset yang akan di ujicoba pada Algoritma C4.5 menggunakan tools Rapidminer. Dataset yang digunakan sebanyak 165 data yang terdiri dari 106 record dengan status Y (Layak) dan 59 record dengan status T (Tidak Layak). Dataset kredit anggota dapat dilihat pada gambar berikut:

	A	B	C	D	E	F	G	H
1	Masa Anggota/bln	Jenis Pinjaman	Nominal Pengajuan	Jangka Waktu	Gaji Pokok	Payroll	Over Limit	CLASS
2	36	Kredit Jangka Pendek	5,000,000	10	12,542,100	4,342,880	35%	Y
3	21	Kredit Jangka Pendek	5,000,000	10	13,038,500	4,589,700	35%	Y
4	6	Kredit Jangka Pendek	5,000,000	10	9,552,300	0	0%	T
5	44	Kredit Jangka Pendek	5,000,000	6	8,953,200	2,181,240	24%	Y
6	65	Kredit Jangka Pendek	5,000,000	8	11,768,600	2,342,100	20%	Y
7	23	Koperasi Karyawan	3,000,000	10	9,858,400	664,900	6.74%	Y
8	7	Koperasi Karyawan	3,000,000	10	9,195,400	0	0%	Y
9	76	Koperasi Karyawan	3,000,000	10	17,520,100	3,454,200	20%	Y
10	43	Koperasi Karyawan	3,000,000	10	9,982,000	5,694,550	57%	T
11	24	Emergency	1,000,000	4	9,652,100	3,543,900	37%	Y
12	6	Koperasi Karyawan	2,000,000	5	10,985,400	0	0%	T
13	34	Koperasi Karyawan	3,000,000	6	13,988,200	4,564,200	33%	Y
14	24	Koperasi Karyawan	2,000,000	5	10,865,900	3,198,600	29%	Y
15	26	Kredit Jangka Pendek	5,000,000	10	12,215,500	4,305,880	35%	Y
16	6	Kredit Jangka Pendek	5,000,000	12	13,040,500	0	0%	T
17	43	Kredit Jangka Pendek	5,000,000	10	9,151,400	3,145,100	34%	Y
18	6	Kredit Jangka Pendek	5,000,000	10	8,944,200	0	0%	T
19	6	Kredit Jangka Pendek	5,000,000	10	9,768,600	0	0%	T
20	6	Koperasi Karyawan	3,000,000	10	9,876,400	0	0.00%	T
21	10	Koperasi Karyawan	3,000,000	10	9,652,800	2,433,100	25%	Y
22	36	Koperasi Karyawan	3,000,000	10	9,520,100	3,368,743	35%	Y

Sumber : Kopkar PT. HPPM

Gambar 2. Dataset kredit anggota

Data tersebut terdapat 8 atribut, 7 diantaranya adalah atribut perbandingan dan 1 atribut hasil untuk menentukan Keputusan. Untuk pengertian masing-masing atribut dapat dilihat pada berikut:

Tabel 1. Atribut data

No	Atribut	Keterangan
1	Masa Anggota	Merupakan lamanya karyawan menjadi anggota koperasi
2	Jenis Pinjaman	Merupakan jenis pinjaman yang akan di ajukan anggota
3	Nominal Pengajuan	Merupakan jumlah pengajuan kredit yang di ajukan anggota
4	Jangka Waktu	Merupakan jangka waktu lamanya pinjaman yang di ajukan anggota
5	Gaji Pokok	Merupakan upah karyawan selama satu bulan terakhir
6	Payroll	Merupakan angsuran perbulan yang harus di bayar
7	Over Limit	Merupakan batas pinjaman perbulan dari upah (gaji pokok)
8	Status	Merupakan kesimpulan analisa mengenai layak atau tidaknya anggota mendapat pinjaman

Sumber : Penelitian, 2024

Tahap selanjutnya dataset yang diakumulasikan akan di klasifikasikan berdasarkan atribut yang telah ditentukan untuk dilakukan pengkajian Algoritma C4.5 dengan menggunakan Rapidminer. Berikut Jumlah atribut dan nilainya:

Tabel 2. Nilai Atribut

No	Atribut	Nilai
1	Masa Anggota / bln	1-6 bln >6 bln

		Kredit Jangka Pendek
2	Jenis Pinjaman	Koperasi Karyawan <i>Emergency</i>
3	Nominal Pengajuan	1-5 Jt >5 Jt
4	Jangka Waktu	1-10 bln >10 bln
5	Gaji Pokok	5-10 Jt >10 Jt
6	<i>Payroll</i>	<5 Jt >5 Jt
7	<i>Over Limit</i>	<40% >40%

Sumber: Penelitian, 2024

3.1.1. Jenis Pinjaman

Koperasi Karyawan pada PT. HPPM ini adalah koperasi simpan pinjam yang menyediakan 3 jenis pinjaman. Berikut Jenis Pinjaman nya :

1. Pinjaman Emergency

Pinjaman Emergency atau pinjaman darurat, pinjaman ini memiliki limit Pinjaman maksimal Rp. 1.000.000,- dan jangka waktu angsuran maksimal 4 bulan.

2. Pinjaman Kopkar

Pinjaman ini biasanya untuk kebutuhan seperti biaya Pendidikan atau kebutuhan lain-lain. Limit pinjaman Kopkar maksimal Rp. 3.000.000,- dan akan dikenakan biaya administrasi sebesar Rp. 100.000,- sehingga anggota akan menerima uang sebesar Rp. 2.900.000,- maksimal jangka waktu angsuran adalah 10 bulan.

3. Kredit Jangka Pendek

Jenis pinjaman ini berfungsi untuk pinjaman pembelian barang atau jasa dengan system reimburse, maksimal limit pengajuan sebesar Rp. 5.000.000,- dan akan dikenakan margin 10% (dari harga beli), maksimal jangka waktu angsuran adalah 10 bulan.

3.1.2. Syarat Pengajuan Pinjaman

Adapun syarat ketentuan pengajuan pinjaman adalah sebagai berikut:

1. Pinjaman hanya diperuntukkan untuk anggota Kopkar PT. HPPM setelah 6 bulan menjadi anggota.
2. Masa angsuran kredit maksimal s.d 10 bulan dan pembayaran dilakukan melalui sistem potong gaji anggota.
3. Besarnya angsuran tidak boleh melebihi 40% dari Gaji Pokok + Tunjangan Tetap.
4. Dalam 1 bulan hanya bisa mengajukan 1 jenis pinjaman.

3.1.3. Alur Pengajuan Kredit

a. Pendaftaran anggota koperasi

Jika karyawan belum menjadi anggota koperasi, maka karyawan harus mengisi formulir untuk menjadi anggota koperasi. Jika sudah menjadi anggota lebih dari 6 bulan, maka baru bisa mengajukan pinjaman.

b. Pengisian formulir

Anggota menghubungi pengelola koperasi untuk mendapatkan formulir pengajuan pinjaman, kemudian anggota mengisi formulir dengan lengkap dan jelas sesuai dengan format pada formulir.

c. Pengajuan dokumen

Setelah mengisi formulir, dan lengkapi dokumen pendukung seperti slip gaji asli 1 bulan terakhir, copy buku rekening tabungan, copy id card/KTP. Kemudian serahkan kepada pengelola koperasi.

d. Verifikasi dan evaluasi

Pengelola koperasi akan melakukan pengecekan perihal pinjamannya untuk apa, jangka waktu cicilannya berapa lama dan nominal pinjamannya berapa disesuaikan juga dengan cicilan perbulannya yang tidak boleh melebihi 40% dari gaji pokok.

e. Keputusan pinjaman

Setelah proses verifikasi dan evaluasi selesai, pengelola koperasi akan memberikan keputusan mengenai pengajuan pinjamannya disetujui atau ditolak. Setelah pengelola koperasi memberi keputusan disetujui, maka pengelola koperasi akan memberikan laporan nya kepada bendahara.

3.2. Penerapan Metode

Algoritma C4.5 digunakan dalam penerapan data mining ini dengan tujuan untuk membuat keputusan layak atau tidak nya anggota Koperasi Karyawan PT. HPPM mendapat pinjaman. Algoritma C4.5 merupakan sebuah algoritma yang digunakan untuk membangun decision tree (pohon keputusan). Pohon Keputusan akan memebentuk akar pohon, selanjutnya data dibedakan sesuai dengan atribut yang sama untuk dibentuk daun atau cabang. Dalam memilih sebuah atribut menjadi akar, dilakukan perhitungan nilai dari atribut yang ada. Nilai gain yang paling tinggi dijadikan root pada pohon Keputusan. Sebelum mencari nilai gain, terlebih dahulu menentukan nilai entropy. Hasil seluruh perhitungan bisa dilihat pada tabel berikut:

Tabel 3. Perhitungan Entropy dan Gain Total

Node	Atribut	jml kasus (s)	Ya (S1)	Tidak (S2)	Entropy	Gain
		165	106	59	0.9406525	
1	Masa Anggota / bln					0.1363727
	1 - 6 bln	14	0	14	0	
	>6 bln	151	106	45	0.8788487	
2	Jenis Pinjaman					0.0168061
	Kredit Jangka Pendek	62	34	28	0.9932338	
	Koperasi Karyawan	85	59	26	0.888363	
	Emergency	18	13	5	0.8524052	
3	Nominal Pengajuan					0.0461979
	1-5 Jt	160	106	54	0.9224063	
	>5 Jt	5	0	5	0	
4	Jangka Waktu					0.2620687
	1-10 bln	140	106	34	0.7997594	
	>10 bln	25	0	25	0	
5	Gaji Pokok					0.0051257
	5-10 Jt	100	61	39	0.9647995	
	> 10 Jt	65	45	20	0.8904916	
6	Payroll					0.0363294
	<5 Jt	153	103	50	0.9116208	
	>5 Jt	12	3	9	0.8112781	
7	Over Limit					0.1939015
	<40%	20	0	20	0	
	>40%	145	105	40	0.8497511	

Sumber : Data Olahan, 2024

Berikut adalah cara menghitung nilai entropy dan nilai gain secara manual dengan menggunakan persamaan berikut :

1. Nilai Entropy

a. Entropy Total = $((-106/165) * \text{LOG}_2 (106/165) + (-59/165) * \text{LOG}_2 (59/165))$

Entropy Total = 0.9406525

b. Entropy Masa Anggota

(1-6 bln) = $((-0/14 * \text{LOG}_2 (0/14) + (-14/14) * \text{LOG}_2 (14/14)) = 0$

(>6 bln) = $((-106/151 * \text{LOG}_2 (106/151) + (-45/151) * \text{LOG}_2 (45/151) = 0.878848729$

c. Entropy Jenis Pinjaman

(Kredit Jangka Pendek) = $((-34/62 * \text{LOG}_2 (34/62) + (-28/62) * \text{LOG}_2 (28/62) = 0.99323382$

(Koperasi Karyawan) = $((-59/85 * \text{LOG}_2 (59/85) + (-26/85) * \text{LOG}_2 (26/85) = 0.888363023$

(Emergency) = $((-13/18 * \text{LOG}_2 (13/18) + (-5/18) * \text{LOG}_2 (5/18) = 0.852405179$

d. Entropy Nominal Pengajian

$$(1-5 Jt) = ((-106/160 * \text{LOG}_2 (106/160) + (-54/160) * \text{LOG}_2 (54/160)) = 0.922406262$$

$$(>5 Jt) = ((-0/5 * \text{LOG}_2 (0/5) + (-5/5) * \text{LOG}_2 (5/5)) = 0$$

e. Entropy Jangka Waktu

$$(1-10 bln) = ((-106/140 * \text{LOG}_2 (106/140) + (-34/140) * \text{LOG}_2 (34/140)) = 0.799759411$$

$$(>10 bln) = ((-0/25 * \text{LOG}_2 (0/25) + (-25/25) * \text{LOG}_2 (25/25)) = 0$$

f. Entropy Gaji Pokok

$$(5-10 Jt) = ((-61/100 * \text{LOG}_2 (61/100) + (-39/100) * \text{LOG}_2 (39/100)) = 0.964799549$$

$$(>10 Jt) = ((-45/65 * \text{LOG}_2 (45/65) + (-20/65) * \text{LOG}_2 (20/65)) = 0.89049164$$

g. Entropy Payroll

$$(<5 Jt) = ((-103/153 * \text{LOG}_2 (103/153) + (-50/153) * \text{LOG}_2 (50/153)) = 0.911620759$$

$$(>5 Jt) = ((-3/12 * \text{LOG}_2 (3/12) + (-9/12) * \text{LOG}_2 (9/12)) = 0.811278124$$

h. Entropy Over Limit

$$(<40\%) = ((-0/20 * \text{LOG}_2 (0/20) + (-20/20) * \text{LOG}_2 (20/20)) = 0$$

$$(>40\%) = ((-105/145 * \text{LOG}_2 (105/145) + (-40/145) * \text{LOG}_2 (40/145)) = 0.849751137$$

2. Nilai Gain

$$\text{Gain Total (Masa Anggota)} = (0.9406525) - ((14/165) * 0) - ((151/165) * 0.8788487) = 0.1363727$$

$$\text{Gain Total (Jenis Pinjaman)} = (0.9406525) - ((62/165) * 0.9932338) - ((85/165) * 0.888363) - ((18/165) * 0.8524052) = 0.0168061$$

$$\text{Gain Total (Nominal Pengajian)} = (0.9406525) - ((160/165) * 0.9224063) - ((5/165) * 0) = 0.0461979$$

$$\text{Gain Total (Jangka Waktu)} = (0.9406525) - ((140/165) * 0.7997594) - ((25/165) * 0) = 0.2620687$$

$$\text{Gain Total (Gaji Pokok)} = (0.9406525) - ((100/165) * 0.9647995) - ((65/165) * 0.8904916) = 0.0051257$$

$$\text{Gain Total (Payroll)} = (0.9406525) - ((153/165) * 0.9116208) - ((12/165) * 0.8112781) = 0.0363294$$

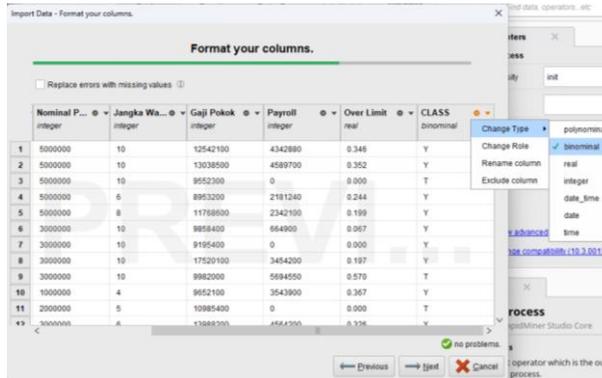
$$\text{Gain Total (Over Limit)} = (0.9406525) - ((20/165) * 0) - ((145/165) * 0.8497511) = 0.1939015$$

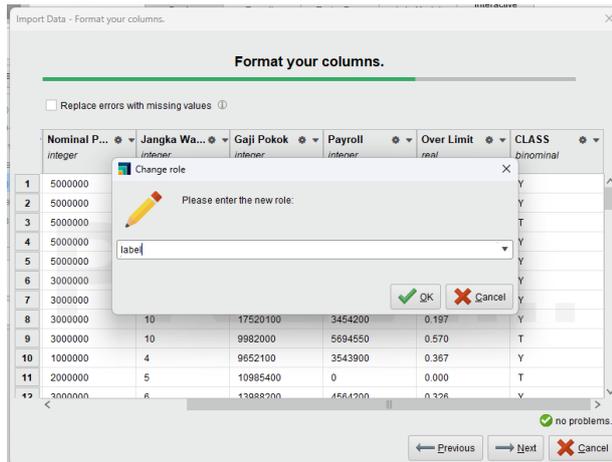
Berdasarkan hasil perhitungan entropy dan gain, bahwa atribut Jangka Waktu memiliki nilai gain tertinggi yaitu 0.2620687. Oleh karena itu, jangka waktu menjadi akar atau node pertama dari pohon Keputusan yang terbentuk.

3.3. Pengolahan Data

3.3.1. Tahap Pre-Processing Data

Tahap proses pengolahan data di aplikasi RapidMiner pada bagian class untuk type nya diganti ke binominal karna memiliki dua nilai dan role nya diubah menjadi label. Pengolahan data akan secara otomatis melakukan pre-processing data yaitu seperti mendeteksi data yang kosong, isi kolom yang berbeda dan akan otomatis memberi centang no problems bila data sudah sesuai dan tidak ada masalah.



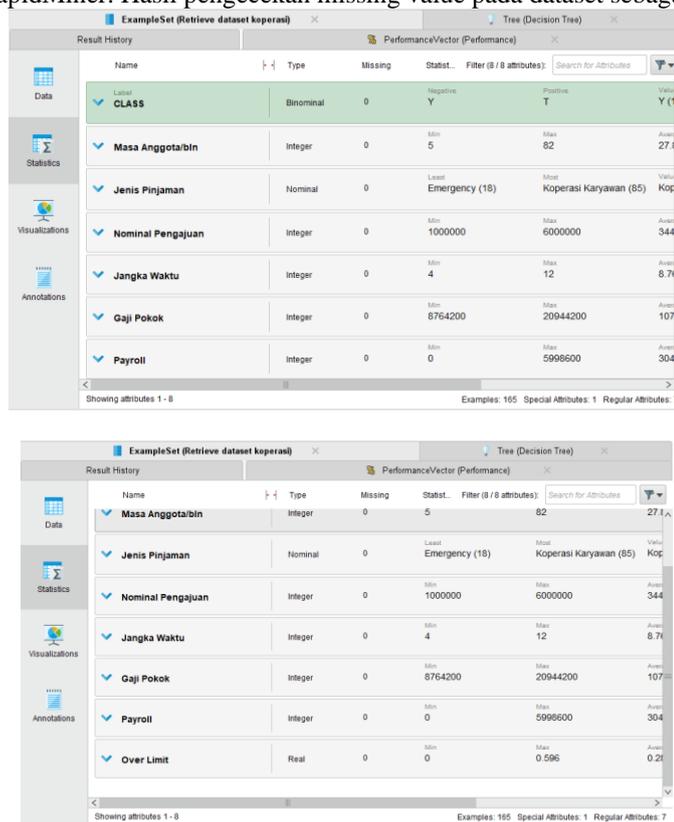


Sumber :Olah Data Rapidminer, 2024

Gambar 3. Proses Pengolahan Data pada Rapidminer

3.3.2. Pengecekan Missing Value

Setelah dataset sudah didapatkan, pastikan tidak terdapat missing value didalamnya untuk mempermudah dalam pengolahan data pada RapidMiner. Hasil pengecekan missing value pada dataset sebagai berikut:

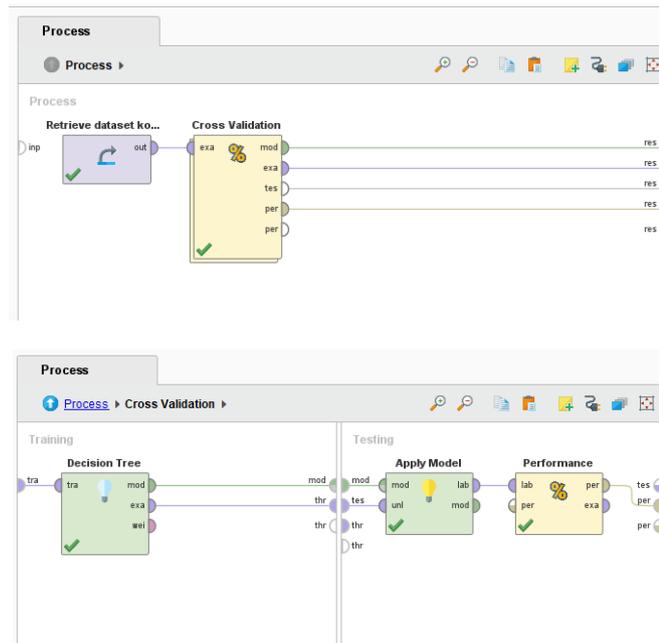


Sumber : Hasil pengecekan pada Rapidminer

Gambar 4. Hasil Pengecekan Missing Value

3.3.3. Proses Pengujian

Pengujian data dilakukan menggunakan tools Rapidminer, pada tahap dataset memakai operator cross validation dengan menggunakan 10 folds serta memakai shuffled sampling untuk mengetahui accuracy, precision, recall dan kurva ROC.



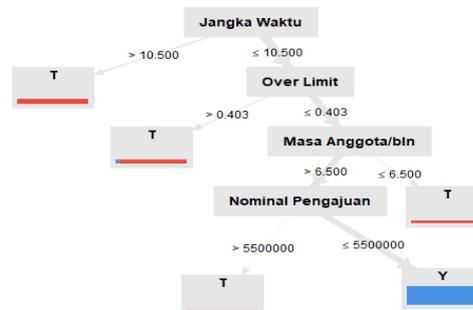
Sumber : Proses pengujian Rapidminer, 2024

Gambar 5. Penggunaan operator Rapidminer

3.4. Analisa Hasil

Hasil dari proses pengujian yang dilakukan dengan tools Rapid miner dapat menghasilkan Decision Tree (pohon keputusan) dan Tabel Confusion Matrix, Tabel tersebut digunakan sebagai salah satu metode evaluasi untuk klasifikasi data, menghitung nilai akurasi, nilai presisinya, nilai recallnya dan Kurva ROC nya.

1. Decision Tree



Sumber: Hasil olahan Rapidminer, 2024

Gambar 6. Decision Tree

Dari hasil Decision Tree diatas terdapat beberapa rules yang dapat dijadikan sebagai referensi dalam klasifikasi Kelayakan pemberian pinjaman pada anggota koperasi. Adapun antara rules diatas berupa satu rules keputusan Y (Layak) dan empat rules keputusan T (Tidak Layak) dapat dijelaskan sebagai berikut:

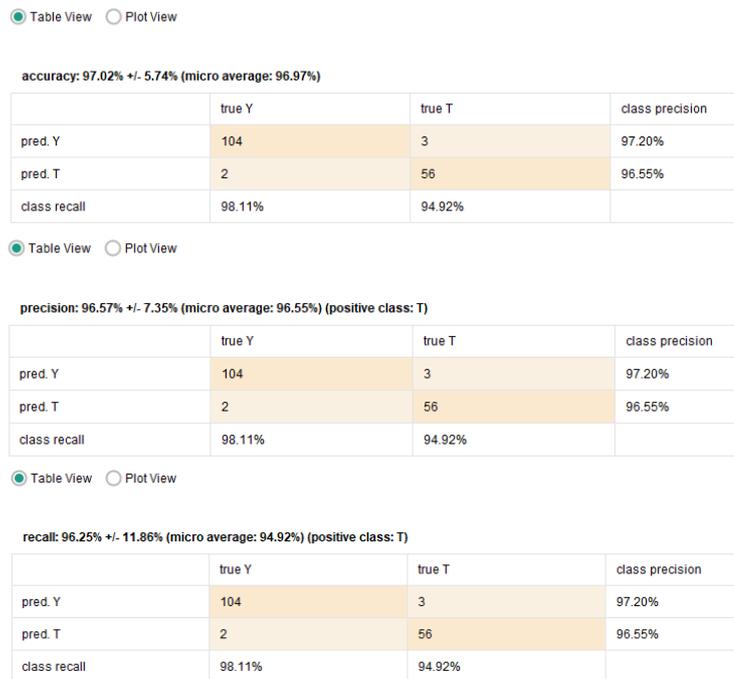
- a. Jika Jangka Waktu >10 bln maka hasil akan menunjukkan T (Tidak Layak) mendapat pinjaman.
- b. Jika Jangka Waktu <10 bln, cek Over Limit >40% maka hasil akan menunjukkan T (Tidak Layak) mendapat pinjaman.
- c. Jika Over Limit <40%, cek Masa Anggota/bln <6 bln maka hasil akan menunjukkan T (Tidak Layak) mendapat pinjaman.
- d. Jika Masa Anggota >6 bln, cek Nominal Pengajuan >5.500.000 maka hasil akan menunjukkan T (Tidak Layak) mendapat pinjaman.
- e. Jika Nominal Pengajuan <5.500.000 maka hasil akan menunjukkan Y (Layak) mendapat pinjaman.

2. Accuracy, Precision, Recall

Accuracy didefinisikan sebagai tingkat kedekatan antara nilai prediksi dengan nilai actual, accuracy yang diperoleh dari hasil pengujian adalah 97,02%.

Precision merupakan tingkat ketepatan antara informasi yang diminta oleh pengguna dengan jawaban yang diberikan oleh sistem dengan data berupa persentase, nilai precision yang diperoleh dari hasil pengujian adalah 96.57% .

Recall adalah tingkat keberhasilan sistem dalam menemukan kembali sebuah informasi. nilai recall yang diperoleh dari hasil pengujian adalah 96.25%. Untuk melihat nilai accuracy, precision dan recall dapat dilihat pada gambar berikut:

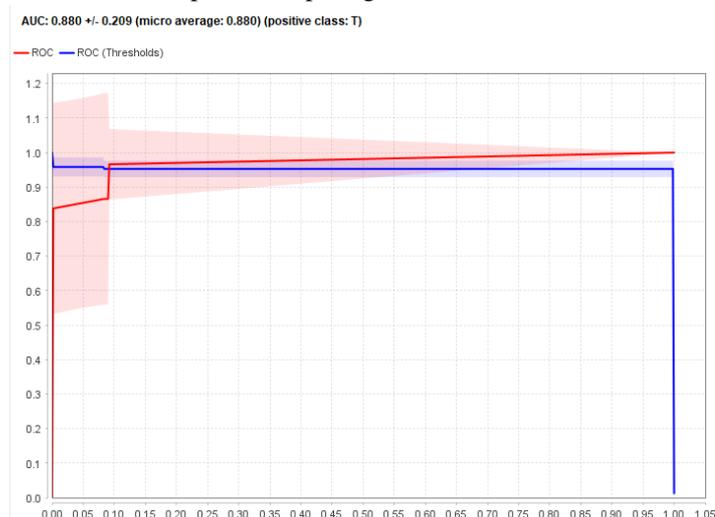


Sumber : Hasil olahan Rapidminer, 2024

Gambar 7. Hasil Nilai 2. Accuracy, Precision, Recall

3. Kurva ROC

Kurva Receiver Operating Characteristic (ROC) digunakan untuk mengekspresikan data confusion matrix. Untuk melihat nilai Kurva ROC dapat dilihat pada gambar berikut ini:



Sumber: Hasil olahan Rapidminer, 2024

Gambar 8. Kurva ROC

Berdasarkan gambar di atas nilai AUC yang dihasilkan sebesar 0.880 yang dapat dikategorikan klasifikasi Baik, dikarenakan memiliki nilai diantara 0.80 – 0.90 maka akurasi memiliki tingkat klasifikasi baik.

5. Confusion Matrix

Confusion matrix disebut sebagai matriks prediksi yang nantinya diperlukan untuk dibandingkan dengan data pada kategori asli. Perhitungan akurasi yang digunakan algoritma menggunakan confusion matrix. Kegunaannya adalah mencari nilai akurasi serta menemukan performa terbaik. Nilai confusion matrix pada Rapidminer menunjukkan hasil dengan tingkat akurasi sebesar 97.02%

```

PerformanceVector:
accuracy: 97.02% +/- 5.74% (micro average: 96.97%)
ConfusionMatrix:
True: Y T
Y: 104 3
T: 2 56
precision: 96.57% +/- 7.35% (micro average: 96.55%) (positive class: T)
ConfusionMatrix:
True: Y T
Y: 104 3
T: 2 56
recall: 96.25% +/- 11.86% (micro average: 94.92%) (positive class: T)
ConfusionMatrix:
True: Y T
Y: 104 3
T: 2 56
AUC (optimistic): 1.000 +/- 0.000 (micro average: 1.000) (positive class: T)
AUC: 0.880 +/- 0.209 (micro average: 0.880) (positive class: T)
AUC (pessimistic): 0.951 +/- 0.118 (micro average: 0.951) (positive class: T)
    
```

Sumber: Hasil olahan Rapidminer, 2024

Gambar 9. Performance Vector

Diketahui jumlah data yang diolah adalah 165 data yang terdiri dari 104 mendapatkan Y Layak dan 56 mendapatkan T Tidak Layak, namun setelah dilakukan pemodelan menghasilkan nilai 104 data diklasifikasikan mendapatkan Y Layak dan sesuai dengan prediksi yang dilakukan dengan metode C4.5, lalu 3 data diprediksi mendapatkan Y Layak tetapi ternyata mendapatkan T Tidak Layak, 2 data di prediksi T Tidak Layak tapi hasilnya mendapatkan Y Layak dan 56 data di prediksi T Tidak Layak dan sesuai dengan yang diprediksikan. Untuk penjelasan gambar di atas, bisa juga di jelaskan pada tabel berikut:

Tabel 4. Confusin Matrix

Performance Vector		
accuracy: 97.02% +/- 5.74% (micro average: 96.97%)		
ConfusionMatrix:		
True:	Y	T
Y:	104	3
T:	2	56
Precision: 96.57% +/- 7.35% (micro average: 96.55%) (positive class: T)		
ConfusionMatrix:		
True:	Y	T
Y:	104	3
T:	2	56
Recall: 96.25% +/- 11.86% (micro average: 94.92%) (positive class: T)		
ConfusionMatrix:		
True:	Y	T
Y:	104	3
T:	2	56
AUC (optimistic): 1.000 +/- 0.000 (micro average: 1.000) (positive class: T)		
AUC: 0.880 +/- 0.209 (micro average: 0.880) (positive class: T)		
AUC (pessimistic): 0.951 +/- 0.118 (micro average: 0.951) (positive class: T)		

Sumber: Penelitian, 2024

Pada tabel diatas merupakan hasil prediksi dan true yang dihasilkan oleh pemodelan algoritma C4.5. dan berikut adalah perhitungan manual untuk ConfusionMatrix

$$\begin{aligned}
 \text{Accuracy} &= (TP + TN) / (TP + FP + FN + TN) = (56+104) / (56+2+3+104) = 0,9696969697 \\
 \text{Precision} &= TP / (TP + FP) = 56 / (56+2) = 0,96551724138 \\
 \text{Recall} &= TP / (TP + FN) = 56 / (56+3) = 0,94915254237 \\
 \text{F1-Score} &= 2 * TP / 2 * (TP + FP + FN) = 2*56/2*(56+2+3) = 0,91803278689
 \end{aligned}$$

Keterangan:

- TP = Jumlah data T diklarifikasi T
- FN = Jjumlah data Y diklarifikasi T
- FP = Jumlah data T diklarifikasi Y
- TN = Jumlah data Y diklarifikasi Y

Berdasarkan hasil akhir pada penerapan metode algoritma C4.5 yang di uji menggunakan tools Rapidminer ini menghasilkan Pohon Keputusan dengan 4 node penentuan dalam memprediksi penentuan kelayakan kredit, yaitu node pertama yang menghasilkan gain tertinggi sebesar 0.2620687 adalah atribut Jangka Waktu, node kedua dengan nilai gain 0.3127118 adalah atribut Over Limit, node ketiga dengan nilai gain 0.2852237 adalah atribut Masa Anggota, dan node keempat dengan nilai gain 0.2298272 adalah atribut nominal pengajuan. Kemudian hasil perhitungan Confusion Matrix ini dapat menghasilkan nilai akurasi, presisi dan recall. Nilai akurasi yang di dapat adalah 97.02%, nilai presisi yang di dapat adalah 96.57% dan nilai recall yang di dapat adalah 96.25%.

4. Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan pada Koperasi Karyawan PT. HPPM mengenai penggunaan klasifikasi Algoritma C4.5 untuk memprediksi kelayakan pemberian kredit, maka dapat disimpulkan sebagai berikut:

1. Klasifikasi data mining dengan menggunakan Algoritma C4.5 terbukti dapat membantu peneliti dan pengelola koperasi dalam penentuan kelayakan pemberian kredit berdasarkan variabel-variabel yang diperlukan yaitu Masa Anggota, Jenis Pinjaman, Nominal Pengajuan, Jangka Waktu, Gaji Pokok, Payroll dan Over Limit.
2. Penentuan kelayakan pinjaman koperasi pada penelitian ini juga berhasil menyederhanakan mekanisme dalam aktivitas analisa kredit dan mampu memberikan status kelayakan dari data yang dimasukkan dengan cukup akurat, dilihat dari hasil pengujian Algoritma C4.5 telah di peroleh Kumpulan aturan pohon Keputusan yang menunjukkan bahwa atribut Jangka Waktu dengan nilai tertinggi dalam memprediksi kelayakan pemberian kredit pada Koperasi Karyawan PT. HPPM.
3. Kelebihan dari penelitian ini adalah presentase keakuratan prediksi relatif besar dibanding penelitian terdahulu. Berdasarkan pengolahan data menggunakan tools rapidminer menghasilkan nilai akurasi yang ideal, yaitu mendapat hasil akurasi sebesar 97.02%, presisi sebesar 96.57%, dan recall sebesar 96.25% dengan AUC sebesar 0.880. Pada penelitian ini juga menghasilkan jumlah prediksi Y sebanyak 104 data dan jumlah prediksi T sebanyak 56 data.
4. Hasil ini membuktikan bahwa metode Algoritma C4.5 dapat digunakan dengan baik dan artinya bahwa rule yang dihasilkan tingkat kebenaran mendekati 100%.

Referensi

- Amril Mutoi Siregar, S. Kom. . M. Kom., & Adam Puspabhuana, S. kom. , M. kom. (2017). *DATA MINING Pengolahan Data Menjadi Informasi dengan RapidMiner* (Aditya Kusuma Putra, Ed.; Pertama). Kekata Publisher.
- Lestari, S., Badrul, M., Studi Sistem Informasi, P., & Nusa Mandiri Jakarta, S. (2020). *ANUGERAH BINTANG CEMERLANG*. 7(1).
- Najib, A., Nurcahyono, D., Pratma, R., & Setiawan, P. (n.d.). *KLASIFIKASI DIAGNOSA PENYAKIT DIABETES MELLITUS (DM) MENGGUNAKAN ALGORITMA C4.4*.
- Prastyadi Wibawa Rahayu, S. Kom. , M. K., I Gede Iwan Sudipa, S. K. . M. C., Suryani, S. Kom. , M. T., Arie Surachman, M. K., Achmad Ridwan, S. Kom. , M. K., I Gede Mahendra Darmawiguna, S. Kom. , M. S., Ir. Muh. Nurtanzis Sutoyo, S. Kom. , M. Cs. , I., Drs. Isnandar Slamet, M. Sc. , Ph. D., Sitti Harlina, SE. , M. K., & I Made Dendi Maysanjaya, S. Pd. , M. E. (2024). *BUKU AJAR DATA MINING* (Efitra, Ed.; Pertama). PT. Sonpedia Publishing Indonesia.
- Raihan M. R., Chrisnanto Y. H., & Ningsih A. K. (2022). *KLASIFIKASI PENENTUAN KELAYAKAN PINJAMAN KOPERASI DENGAN ALGORITMA CART MENGGUNAKAN ALGORITMA ADABOOST*. *Infotech*, 8, 2.
- Rifkie Primartha, & Romi Satria Wahono, Prof. (2021). *Algoritma Machine Learning*. Informatika Bandung.
- Riswanto, I., & Laluma, R. H. (2020). *CLASSIFIER BERBASIS WEB*. *Jurnal*, 5(1). <https://doi.org/10.32897/infotronik.2020.5.1.2>
- Sodikin, E. S. L., Putra, T. W., Ruhayat, M. D., & Yulianto, A. (2021). *Klasifikasi Data Pinjaman Koperasi Menggunakan Algoritma Naive Bayes*. *Paradigma - Jurnal Komputer Dan Informatika*, 23(2). <https://doi.org/10.31294/p.v23i2.11084>
- Suyanto. (2018). *Machine Learning Tingkat Dasar dan Lanjut* (Pertama). Informatika Bandung.