

## PENERAPAN METODE DECISION TREE C4.5 DALAM MEMPREDIKSI KELANCARAN PEMBAYARAN KREDIT DI BPR WAHANA BERSAMA KPUM

Nine Juli, Indra Kelana Jaya✉, Margaretha Yohanna

Teknik Informatika, Universitas Methodist Indonesia, Medan, Indonesia

Email: [indraikjs@gmail.com](mailto:indraikjs@gmail.com)

DOI: <https://doi.org/10.46880/jmika.Vol6No2.pp215-220>

### ABSTRACT

*BPR Wahana Bersama KPUM is a lending company. Delay in credit payments is a problem that occurs in credit companies. Customers who don't pay on time can have a bad impact on their credit history. To assess customer profitability, a system is needed that can predict the smoothness of future credit payments in order to assess whether customers are profitable or not. The Decision Tree C4.5 method is used in this study because the method looks for similarities between classes or groups in the data. In this study, customer data is used, with training data from 2017-2020 and testing data from 2021. The existence of this system can help BPR Wahana Bersama KPUM in predicting the smoothness of credit payments in the future so that there will be no congestion in credit payments by customers, this is proven by the acquisition of the accuracy value using the confusion matrix model reaching 84.18%.*

**Keyword:** Prediction, Credit, C4.5.

### ABSTRAK

BPR Wahana Bersama KPUM merupakan perusahaan pemberi pinjaman. Keterlambatan pembayaran kredit merupakan masalah yang terjadi di perusahaan perkreditan. Nasabah yang tidak membayar tepat waktu dapat berdampak buruk pada riwayat kredit mereka. Untuk menilai profitabilitas nasabah, diperlukan sistem yang dapat memprediksi kelancaran pembayaran kredit di masa mendatang agar dapat menilai nasabah yang menguntungkan atau tidak. Metode Decision Tree C4.5 digunakan pada penelitian ini dikarenakan metode tersebut mencari kesamaan antar kelas atau kelompok dalam data. Pada penelitian ini, digunakan data nasabah, dengan data training dari tahun 2017-2020 dan data testing dari tahun 2021. Adanya sistem ini dapat membantu pihak BPR Wahana Bersama KPUM dalam memprediksi kelancaran pembayaran kredit kedepannya agar tidak terjadi kemacetan pembayaran kredit oleh nasabah, hal ini terbukti dengan perolehan nilai akurasi menggunakan model confusion matrix mencapai 84.18%.

**Kata Kunci:** Prediksi, Kredit, C4.5.

### PENDAHULUAN

BPR Wahana Bersama KPUM merupakan lembaga pemberian kredit. Peminjaman kredit dilakukan untuk pembiayaan nasabah, kebutuhan sehari-hari atau pengembangan bisnis. Proses pinjaman kredit berlangsung paling lama 8 tahun. Namun, dalam hal pembayaran kredit banyak nasabah yang bermasalah dalam keterlambatan pembayaran kredit. Informasi dari nasabah hanya berasal dari hasil survei yang menentukan nasabah mana yang memiliki akses dalam peminjaman kredit. Hal tersebut yang menyebabkan terjadinya nasabah menunggak pembayaran. Oleh karena itu, untuk meminimalkan kerugian memerlukan pertimbangan banyak faktor.

Sebelum memberikan kredit, diperlukan suatu sistem untuk memprediksi nasabah yang akan *default* karena keterlambatan pembayaran. Hal ini

memungkinkan untuk keputusan yang akan dibuat mengenai nasabah ini. Metode yang dapat digunakan dalam memprediksi kelancaran pembayaran kredit adalah *Decision Tree C4.5*.

Gartner Group menggambarkan *data mining* sebagai proses menyeleksi sejumlah data pada penyimpanan dengan metode statistik dan matematika yang akan menghasilkan informasi baru dengan arti dan pola (Mardi, 2017). Dalam penerapan *data mining*, pohon keputusan (*decision tree*) merupakan suatu cara untuk memprediksi/mengklasifikasi yang sangat kuat. Pohon keputusan merupakan struktur yang dipakai dalam pembagian kumpulan data besar menjadi kumpulan *record* yang lebih kecil dengan penerapan *rule*. C4.5 adalah salah satu algoritma pohon keputusan yang dapat diterapkan untuk mengklasifikasikan atau

mengelompokkan record dengan sifat prediktif (Setio, Saputro, & Winarno, 2020).

**TINJAUAN PUSTAKA**

**Decision Tree C4.5**

Salah satu data mining klasifikasi yaitu algoritma *decision tree*. Algoritma *decision tree* menggunakan data latih yang ada pada *database* untuk membangun sebuah pohon keputusan. Algoritma pohon keputusan terdiri dari kumpulan simpul (*node*) yang dihubungkan oleh cabang-cabang yang bergerak ke bawah dari *node* akar (*root*) dan berakhir di *node* daun (*leaf*). *Leaf node* tidak dapat dipecah lagi dan merupakan hasil yang diprediksi berdasarkan data yang diuji oleh *node* yang diberikan (Huda, 2019; Suntoro, 2019).

C4.5 adalah salah satu jenis algoritma pohon keputusan yang mengklasifikasikan dataset dengan mengelompokkannya ke dalam kelompok tertentu. Tujuan dari algoritma C4.5 adalah untuk mengklasifikasikan kumpulan data, sehingga dapat dikelompokkan ke dalam kelas-kelas tertentu setelah diproses (Iriadi & Nuraeni, 2016)

Algoritma C4.5 menggunakan *entropy*, *information gain*, *split info* dan *gain ratio* untuk memilih atribut yang akan dijadikan *node* (Hermawanti, Asriyanik, & Sunarto, 2019). Langkah-langkah algoritma C4.5 untuk membangun pohon keputusan adalah sebagai berikut:

1. Siapkan dataset
2. Pilih Variabel/Atribut sebagai akar
3. Bagi Kasus dalam cabang
4. Ulangi proses untuk masing-masing cabang sampai semua kasus pada cabang memiliki kelas yang sama.

Penentuan variabel/atribut yang menjadi akar dari pohon keputusan menggunakan beberapa rumus dibawah ini:

1. Hitung Nilai *Entropy*  

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

2. Hitung Nilai *Information Gain*  

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} \times Entropy(S_i) \quad \dots\dots (2)$$

3. Hitung Nilai *Split Info*  

$$SplitInfo_A(D) = - \sum_{j=1}^v \frac{|D_j|}{|D|} \times \log_2 \left( \frac{|D_j|}{|D|} \right) \quad \dots\dots (3)$$

4. Hitung Nilai *Gain Ratio*  

$$GainRatio = \frac{Gain(S,A)}{SplitInfo_A(D)} \quad \dots\dots (4)$$

**Confusion Matrix**

Metode *confusion matrix* menggunakan tabel matriks seperti pada Tabel 1, dataset yang terdiri dari 2 kelas klasifikasi, kelas pertama positif dan kelas kedua negatif (Mardiana, 2018).

**Tabel 1.** Model *Confusion Matrix*

Correct Classification	Classified as	
	+	-
+	True Positives	False Negatives
-	False Positives	True Negatives

*True Positives* yaitu *record* positif yang diklasifikasikan positif, *false positives* yaitu *record* negatif yang diklasifikasikan positif, *false negatives* yaitu *record* positif yang diklasifikasikan negatif, *true negatives* yaitu *record* negatif yang diklasifikasikan negatif.

Untuk menghitung *Accuracy* dan *Error* digunakan persamaan di bawah ini :

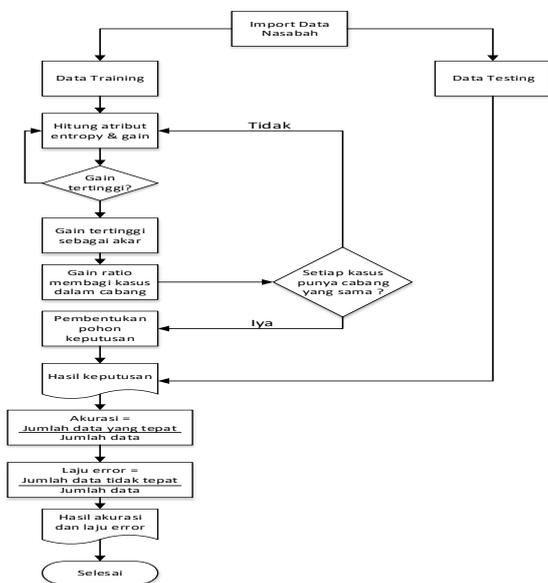
$$Accuracy = \frac{True\ Positives + True\ Negatives}{Total\ Data} \times 100\% \quad \dots\dots (5)$$

$$Error = \frac{False\ Positives + False\ Negatives}{Total\ Data} \times 100\% \quad \dots\dots (6)$$

**METODE PENELITIAN**

**Analisis Sistem**

Proses pada sistem ada 2 bagian, yaitu data *training* dan data *testing*. Tahapan-tahapan yang dilakukan pada sistem dapat dilihat pada gambar 1.



**Gambar 1.** Analisis Sistem

**Analisis Data**

Data yang digunakan yaitu data nasabah dari

tahun 2017-2021 yang didapat dari pihak BPR Wahana Bersama KPUM, dapat dilihat pada Tabel 2.

**Tabel 2.** Data Nasabah

Nama	Status	Tanggung	Pekerjaan	Penghasilan	Pinjaman	Jangka (Bulan)	Bunga	Jaminan	Kelas
TS	M	1	Wiraswasta	37.200.000	7.216.000	22	Flat	Rumah	Macet
MG	M	3	Wiraswasta	113.400.000	110.000.000	60	Flat	Mobil	Macet
BS	M	6	Wiraswasta	48.000.000	110.000.000	60	Flat	Mobil	Lancar
BHS	M	2	Wiraswasta	180.000.000	52.778.000	36	Flat	Mobil	Lancar
Sj	M	4	Wiraswasta	131.850.000	9.247.000	26	Flat	Rumah	Macet
Wh	M	2	Wiraswasta	72.240.000	11.785.000	28	Flat	Rumah	Macet
SS	M	2	Wiraswasta	72.000.000	13.950.000	24	Flat	Rumah	Lancar
MTS	J	1	Pertanian	250.800.000	30.282.000	36	Sidiing	Rumah	Lancar
Nt	M	1	Wiraswasta	54.000.000	65.000.000	60	Flat	Rumah	Macet
BLJP	M	1	Wiraswasta	75.600.000	60.000.000	60	Flat	Rumah	Lancar

Berdasarkan data yang telah dikumpulkan diperoleh kriteria/atribut yang akan digunakan untuk perkiraan dan peninjauan, dapat dilihat pada Tabel 3.

**Tabel 3.** Kriteria

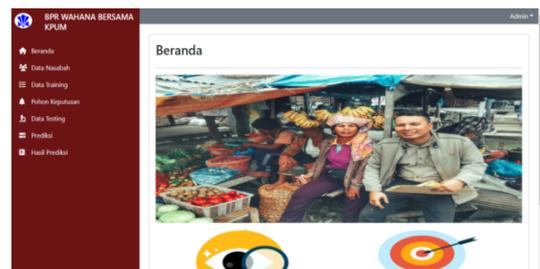
Kriteria	Parameter
Status Pernikahan	Menikah Belum Menikah Janda/Duda
Tanggung	0 1 – 3 orang > 3 orang
Pekerjaan	PNS/BUMN Wiraswasta Petanian/Peternakan
Penghasilan (per tahun)	< 50 Juta 50 – 500 Juta > 500 Juta
Pinjaman	< 50 Juta 50 – 525 Juta 526 – 1000 Juta
Jangka	< 33 Bulan 33 – 64 Bulan 65 – 96 Bulan
Bunga	Flat Persekot Sliding
Jaminan	Rumah Mobil

**HASIL DAN PEMBAHASAN**

Sistem akan menampilkan setiap halaman menu yang digunakan dalam memprediksi kelancaran pembayaran kredit di BPR Wahana Bersama KPUM. Tampilan halaman sistem adalah sebagai berikut:

**Halaman Beranda**

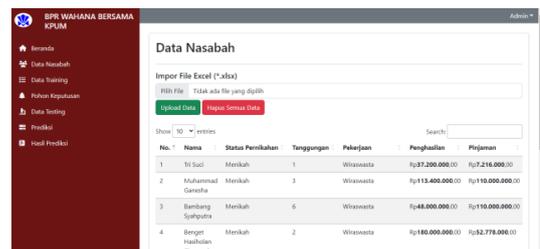
Halaman ini menampilkan profil perusahaan, dapat dilihat pada Gambar 2.



**Gambar 2.** Halaman Beranda

**Halaman Data Nasabah**

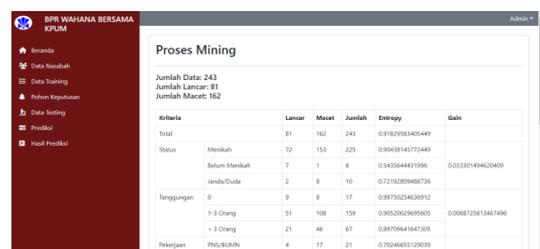
Halaman ini untuk mengupload data nasabah dalam bentuk excel untuk proses data *training*, dapat dilihat pada Gambar 3.



**Gambar 3.** Halaman Data Nasabah

**Halaman Data Training**

Pada halaman ini akan melakukan proses algoritma C4.5 terhadap data nasabah, dapat dilihat pada Gambar 4.



**Gambar 4.** Halaman Data Training

**Halaman Pohon Keputusan**

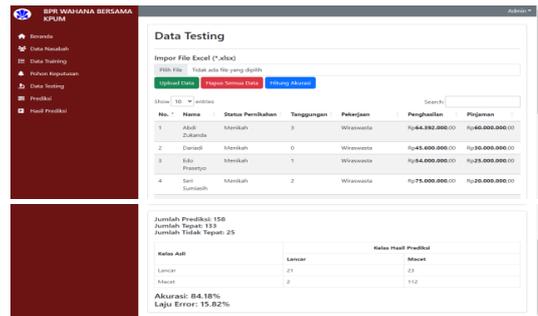
Halaman ini menampilkan hasil dari pohon keputusan yang telah diperoleh, dapat dilihat pada Gambar 5.



**Gambar 5.** Halaman Pohon Keputusan

**Halaman Data Testing**

Halaman ini untuk mengupload data *testing* dalam bentuk excel. Pada halaman ini akan menampilkan hasil kelas prediksi dan akurasi sistem, dapat dilihat pada Gambar 6.



**Gambar 6.** Halaman Data *Testing*

Pada penelitian ada 2 kelas klasifikasi, kelas “Lancar” dan “Macet”. Data *training* sejumlah 243 data dengan 81 data lancar dan 162 data macet. Hasil penerapan algoritma c4.5 iterasi ke-1 dapat dilihat pada Tabel 4.

**Tabel 4.** Penerapan C4.5

Kriteria		Lancar	Macet	Jumlah	Entropy	Gain
<b>Total</b>		81	162	243	0,918	0,033
<b>Status</b>	Menikah	72	153	225	0,904	
	Belum Menikah	7	1	8	0,544	
	Janda/Duda	2	8	10	0,722	
<b>Tanggungan</b>	0	9	8	17	0,998	0,009
	1 – 3 orang	51	108	159	0,905	
	> 3 orang	21	46	67	0,897	
<b>Pekerjaan</b>	PNS / BUMN	4	17	21	0,702	0,013
	Wiraswasta	67	135	202	0,917	
	Pertanian/Peternakan	10	10	20	1,000	
<b>Penghasilan</b>	< 50 jt	29	41	70	0,979	0,009
	50 – 500 jt	51	120	171	0,879	
	> 500 jt	1	1	2	1,000	
<b>Pinjaman</b>	< 50 jt	53	121	174	0,887	0,007
	50 – 525 jt	28	41	69	0,974	
	526 – 1000 jt	0	0	0	0,000	
<b>Jangka</b>	< 33 bulan	13	20	33	0,967	0,057
	33 – 64 bulan	60	142	202	0,878	
	65 – 96 bulan	8	0	8	0,000	
<b>Bunga</b>	Flat	55	130	185	0,878	0,039
	Persekot	5	18	23	0,755	
	Sliding	21	14	35	0,971	
<b>Jaminan</b>	Tanah & Rumah	74	158	232	0,903	0,013
	Mobil	7	4	11	0,946	

Pada iterasi ke-1 diperoleh *gain* terbesar adalah 0.057 yaitu kriteria “Jangka”. Maka **jangka** merupakan akar pohon keputusan (*root*). Dilakukan proses perhitungan untuk *node* berikutnya sampai proses

perhitungan berakhir dengan setiap cabang pohon keputusan sudah memiliki kelas yang sama. Pohon keputusan yang akan menjadi aturan-aturan keputusan (*rule*) untuk memprediksi

**Pengujian Sistem**

Data *testing* yang digunakan adalah data nasabah tahun 2021 yaitu sebanyak 158 data, dengan

kelas lancar 44 data dan kelas macet 114 data, dapat dilihat pada Tabel 5.

Tabel 5. Data Testing

Nama	Status	Tanggungan	Pekerjaan	Penghasilan	Pinjaman	Jangka (Bulan)	Bunga	Jaminan	Kelas Asli	Kelas Hasil
AZ	M	3	Wiraswasta	64.392.000	60.000.000	36	Flat	Rumah	Lancar	Macet
Dr	M	0	Wiraswasta	45.600.000	30.000.000	36	Flat	Rumah	Macet	Macet
EP	M	1	Wiraswasta	54.0000.00	25.000.000	36	Flat	Rumah	Macet	Macet
SS	M	2	Wiraswasta	75.000.000	20.000.000	36	Sliding	Rumah	Macet	Macet
SIPA	M	5	Wiraswasta	87.564.000	80.000.000	48	Flat	Rumah	Macet	Macet
MNH	M	2	Wiraswasta	68.400.000	70.000.000	60	Sliding	Rumah	Lancar	Macet
IK	M	2	Wiraswasta	60.000.000	57.148.000	48	Persekot	Rumah	Macet	Macet
MS	M	4	Pertanian	63.000.000	15.000.000	24	Persekot	Rumah	Lancar	Lancar
BN	M	5	Wiraswasta	107.724.000	10.000.000	24	Flat	Rumah	Macet	Macet
BS	BN	3	PNS/BUMN	22.000.000	9.000.000	60	Flat	Rumah	Macet	Macet

Pengujian data atas nama AZ dengan data Status Pernikahan = Menikah, Tanggungan = 3, Pekerjaan = Wiraswasta, Penghasilan = 64.392.000, Pinjaman = 60.000.000, Jangka = 36 bulan, Bunga = Flat, Jaminan = Rumah, dan kelas asli = Lancar diproses dengan aturan : “IF Jangka = 33-64 bulan, Status Pernikahan = Menikah, Pekerjaan = Wiraswasta, Bunga = Flat, Penghasilan = 50-500 jt, Tanggungan = 1-3 orang, Pinjaman = 50-525 jt, Jaminan = Rumah maka Kelas = ”Macet”. Dengan kelas asli “Lancar” dan kelas hasil “Macet” maka prediksi **salah**.

**Akurasi**

Penelitian ini menggunakan *Confusion Matrix* untuk menguji kinerja sistem, dengan data *testing* sebanyak 158 data yang akan menghasilkan nilai akurasi dan error, dapat dilihat pada Tabel 6.

Tabel 6. Akurasi

Kelas Asli	Kelas Hasil Prediksi	
	Lancar	Macet
Lancar	21	23
Macet	2	112

Efektivitas sistem dievaluasi dengan menghitung nilai TP, TN, FP, dan FN, dan diperoleh nilai akurasi dan error sebagai berikut:

$$Akurasi = \frac{TP+TN}{Total\ Data} \times 100\% = \frac{21+112}{158} \times 100\% = \frac{133}{158} \times 100\% = 84,18\%$$

$$Error = \frac{FP+FN}{Total\ Data} \times 100\% = \frac{2+23}{158} \times 100\% = \frac{25}{158} \times 100\% = 15,82\%$$

Dari hasil evaluasi terbukti bahwa algoritma C4.5 memiliki **84,18%** dalam menentukan kelancaran

pembayaran kredit. Ini berarti aturan C4.5 dapat memudahkan dalam memprediksi pembayaran kredit.

**KESIMPULAN**

Dari perancangan sistem prediksi kelancaran pembayaran kredit dengan menggunakan metode decision tree c4.5 dapat disimpulkan bahwa :

1. Berdasarkan 243 data training, diperoleh 81 nasabah lancar dalam pembayaran kredit dan 162 nasabah macet dalam pembayaran kredit.
2. Dari hasil evaluasi menggunakan confusion matrix, dengan menggunakan 158 data testing dan menggunakan delapan variabel/atribut menghasilkan akurasi sebesar 84,18%.
3. Aturan Aturan keputusan (rules) yang dihasilkan pada sistem prediksi kelancaran pembayaran kredit menggunakan algoritma c4.5 mampu menghasilkan keputusan berdasarkan kriteria yang digunakan.

**DAFTAR PUSTAKA**

Hermawanti, S. N., Asriyanik, & Sunarto, A. A. (2019). Implementasi Algoritma C4. 5 Untuk Prediksi Kelulusan Tepat Waktu (Studi Kasus: Program Studi Teknik Informatika). *SANTIKA Is a Scientific Journal of Science and Technology*, 9(1), 853–864.

Huda, M. (2019). *Algoritma Data Mining: Analisis Data Dengan Komputer*. Bogor: Bisakimia.com.

Iriadi, N., & Nuraeni, N. (2016). Kajian Penerapan Metode Klasifikasi Data Mining Algoritma C4.5 untuk Prediksi Kelayakan Kredit pada Bank Mayapada Jakarta. *Jurnal Teknik Komputer AMIK BSI*, 2(1), 132–137.

Mardi, Y. (2017). Data Mining: Klasifikasi Menggunakan Algoritma C4.5. *Edik Informatika*, 2(2), 213–219. <https://doi.org/10.22202/ei.2016.v2i2.1465>

Mardiana, T. (2018). Optimasi Naïve Bayes Dengan Particle Swarm Optimization Dan Stratified

- Untuk Prediksi Kredit Macet Pada Koperasi.  
*Jurnal Riset Informatika*, 1(1), 43–50.  
<https://doi.org/10.34288/jri.v1i1.13>
- Setio, P. B. N., Saputro, D. R. S., & Winarno, B.  
(2020). Klasifikasi Dengan Pohon Keputusan Berbasis Algoritme C4.5. In *PRISMA 3: Prosiding Seminar Nasional Matematika* (pp. 64–71).
- Suntoro, J. (2019). *DATA MINING: Algoritma dan Implementasi dengan Pemrograman PHP*. Jakarta: Elex Media Komputindo.