

PREDIKSI KELAYAKAN PEMBERIAN KREDIT PADA NASABAH DI KSP BONA MANDIRI JAYA MENGGUNAKAN ALGORITMA DATA MINING C4.5

Nancy Veronica Br Gultom

Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Medan, Medan, Indonesia

e-mail : nancy17.4223230034@mhs.unimed.ac.id*

Suvriadi Panggabean

Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Medan, Medan, Indonesia

e-mail : suvriadi@unimed.ac.id

Abstrak

Penelitian ini bertujuan untuk membangun model klasifikasi dalam memprediksi kelayakan kredit nasabah ulang di KSP Bona Mandiri Jaya dengan menggunakan algoritma C4.5 dalam data mining. Data yang digunakan berjumlah 114 catatan nasabah tahun 2022, yang mencakup atribut riwayat pinjaman, jaminan, penghasilan bulanan, dan status kelayakan. Metode penelitian menerapkan teknik klasifikasi pohon keputusan dan dievaluasi menggunakan *confusion matrix* serta validasi split (80% data latih, 20% data uji) dengan bantuan aplikasi Rapid Miner. Hasil penelitian menunjukkan bahwa algoritma C4.5 mampu membentuk pohon keputusan dengan aturan klasifikasi yang mudah dipahami, seperti: *JIKA riwayat pinjaman = lancar DAN penghasilan > 500000 MAKA layak*. Model menunjukkan performa tinggi dengan akurasi sebesar 98,25%, precision 94,12% untuk kelas tidak layak dan 98,97% untuk kelas layak, serta recall masing-masing sebesar 94,12% dan 98,97%. Model prediktif ini diharapkan dapat membantu petugas kredit dalam membuat keputusan pemberian pinjaman yang lebih objektif, akurat, dan konsisten, serta mengurangi risiko gagal bayar dan mendukung pengelolaan keuangan koperasi yang lebih baik.

Kata Kunci: Kelayakan Kredit, Data Mining, Algoritma C4.5, Pohon Keputusan, Rapid Miner.

Abstract

This study aims to build a classification model to predict the credit eligibility of repeat borrowers at KSP Bona Mandiri Jaya using the C4.5 algorithm in data mining. The dataset used consists of 114 borrower records from 2022, which includes attributes such as loan history, collateral, monthly income, and eligibility status. The research method applies a decision tree classification technique and is evaluated using a confusion matrix and split validation (80% training data, 20% testing data) with the help of the RapidMiner application. The results indicate that the C4.5 algorithm can construct a decision tree with interpretable classification rules, such as: IF loan history = smooth AND income > 500000 THEN eligible. The model demonstrated high performance with an accuracy of 98.25%, precision of 94.12% for the ineligible class and 98.97% for the eligible class, and a recall of 94.12% and 98.97%, respectively. This predictive model is expected to assist credit officers in making more objective, accurate, and consistent decisions regarding loan approvals, while also reducing the risk of loan default and supporting better financial management in cooperatives.

Keywords: Credit Eligibility, Data Mining, C4.5 Algorithm, Decision Tree, Rapid Miner.

PENDAHULUAN

Menurut (Sartika Partomo, 2004) dalam meningkatkan pertumbuhan ekonomi Indonesia, pemerintah berupaya mendorong pengembangan koperasi sebagai salah satu cara untuk memperkuat perekonomian Indonesia (Perkasa et al., 2024). Pemerintah mendukung melalui Undang-Undang Nomor 25 Tahun 1992 tentang Perkoperasian (*Undang-Undang Nomor 25 Tahun 1992 Tentang*

Perkoperasian, 1992). Koperasi Simpan Pinjam (KSP) memiliki peran yang signifikan dalam meningkatkan pemberdayaan masyarakat, terutama bagi para pelaku Usaha Mikro, Kecil, dan Menengah (Maulida et al., 2024).

Dalam penyaluran kredit, koperasi simpan pinjam seringkali menghadapi risiko penurunan kualitas kredit dan gagal bayar sehingga menimbulkan kerugian finansial yang disebabkan oleh faktor internal dan eksternal (Endris, 2022).

Oleh karena itu, perlunya evaluasi mendalam penghargaan kredit menggunakan 5C, yaitu karakter, kapasitas, modal, kondisi ekonomi, agunan (Nisak et al., 2024). Selain prinsip 5C dalam memilih semua kredit yang masuk, ada beberapa pertimbangan yang menjadi dalam lolos aplikasi, yaitu kelengkapan dokumen, risiko kredit berdasarkan aspek pendapatan dan riwayat kerja yang besar, memiliki riwayat kredit yang lancer (Christnatis et al., 2021). Dengan melakukan penilaian kelayakan kredit secara cermat, bank atau lembaga keuangan dapat memastikan bahwa calon nasabah memiliki kemampuan dan kemauan untuk memenuhi kewajiban pembayaran kredit sesuai perjanjian demi menjaga stabilitas keuangan kedua belah pihak dan mengurangi risiko kerugian yang tidak diinginkan (Sholihaningtias, 2023).

Menurut (Syahfitri, 2017) untuk mengevaluasi potensi kredit peminjam ulang, analisis kredit menggunakan Data Mining sangat diperlukan. Proses ini melibatkan penggunaan teknologi untuk memeriksa dan menganalisis sejumlah besar data secara otomatis atau semi-otomatis, dengan tujuan menemukan pola atau aturan yang signifikan dalam (Oktafriani et al., 2023). Menurut (Fimawahi & Rouza, 2021) Untuk meningkatkan akurasi penelitian, teknik validasi split yang digunakan adalah validasi Rapid Miner dengan cara membagi data menjadi dua bagian yaitu, data pengujian dan memvalidasi model yang kemudian digunakan untuk memvalidasi model dalam (Puspita et al., 2022).

Berbagai penelitian sebelumnya telah membuktikan bahwa algoritma C4.5 mampu memberikan hasil klasifikasi yang akurat dalam konteks penilaian kelayakan kredit. Penelitian oleh (Sutjiadi et al., 2022) merancang sistem informasi penjualan berbasis web yang terintegrasi dengan fitur penentuan kelayakan kredit menggunakan algoritma C4.5, dan berhasil memperoleh nilai precision sebesar 0,876, recall 0,952, serta F-measure 0,912, menunjukkan performa klasifikasi yang sangat baik. Penelitian serupa oleh (Tsalatsatun Nur Rohmah, 2024) menerapkan algoritma C4.5 pada lembaga keuangan untuk mengidentifikasi risiko kredit dengan dukungan perangkat lunak RapidMiner, dan menghasilkan akurasi prediksi sebesar 97,78% menggunakan 540 data training dan 135 data testing. Penelitian oleh (Wirasena & Warmansyah, 2024) memanfaatkan variabel seperti status rumah, penghasilan, status pernikahan, dan usia untuk memprediksi kelayakan pengajuan kartu kredit Visa, dengan akurasi sebesar 83,33% menggunakan confusion matrix. Penelitian oleh (Gunawan & Hasibuan, 2023) melakukan perbandingan algoritma C4.5 dan Naïve Bayes

dalam penentuan penerima bantuan bedah rumah, dan menunjukkan bahwa C4.5 memberikan hasil klasifikasi yang lebih akurat dan tepat sasaran dalam penyaluran bantuan. Sementara itu, penelitian oleh (Martin & Nilawati, 2021) membandingkan kedua algoritma tersebut dalam menilai pelayanan pengaduan publik di Komnas HAM, dan menemukan bahwa C4.5 unggul dengan akurasi 99,49% serta AUC sebesar 0,998, lebih tinggi dibandingkan Naïve Bayes.

Berdasarkan latar belakang dan tinjauan penelitian sebelumnya, algoritma C4.5 menunjukkan keberhasilan dalam klasifikasi kelayakan kredit dan layanan publik. Namun sebagian besar studi tersebut belum secara khusus mengkaji penerapannya dalam konteks koperasi, terutama pada kasus evaluasi kredit ulang nasabah aktif. Selain itu, masih jarang ditemukan penelitian yang menggabungkan metode C4.5 dengan validasi model berbasis split di Rapid Miner untuk meningkatkan akurasi prediksi. Dengan demikian, penelitian ini difokuskan untuk memprediksi kelayakan pemberian kredit ulang bagi nasabah di KSP Bona Mandiri Jaya menggunakan Algoritma data mining C4.5 dan validasi Rapid Miner split dalam meningkatkan keakuratan prediksi kelayakan pemberian kredit. Diharapkan hasil penelitian ini dapat membantu lembaga koperasi dalam mengurangi risiko gagal bayar dan memastikan proses penilaian kredit yang lebih objektif dan terukur.

KAJIAN TEORI

A. Data Mining

Data mining adalah proses menggali informasi dari himpunan data yang berukuran besar dan kompleks. Tujuan utamanya adalah menemukan pola, keterkaitan, atau pengetahuan yang tidak mudah terlihat secara langsung dalam data, sehingga dapat menghasilkan wawasan yang lebih mendalam dan bermanfaat (Ville, 2001).

Salah satu teknik data mining yang digunakan untuk mengekstrak pola dan pengetahuan dari data yaitu Klasifikasi. Klasifikasi mengacu pada pendekatan nilai variabel target numerik (*Independent variabel*) yang menggunakan variabel prediktor (*Dependent variabel*). Metode klasifikasi mengacu pada pengklasifikasian data menurut kelas kategoris yang telah ditentukan sebelumnya (Muttaqien et al., 2021).

B. Decision Tree dan Algoritma C4.5

Decision Tree adalah salah satu teknik klasifikasi dalam data mining (Firmansyah & Aufany, 2016). Algoritma C4.5 digunakan untuk membangun pohon keputusan. Secara garis besar, pembentukan decision tree dengan algoritma C4.5 dilakukan melalui beberapa tahapan umum berikut (Waluyo et al., 2020) :

1. Menyiapkan data training dari record yang ada sebelumnya
2. Pemilihan akar atribut
3. Pembuatan cabang
4. Membagi kasus
5. Hingga cabang mempunyai kelas yang sama, ulangi lagi prosesnya.

Pemilihan atribut sebagai akar dilakukan berdasarkan atribut yang memiliki nilai gain tertinggi. Untuk memperoleh nilai gain, langkah awal yang perlu dilakukan adalah menghitung nilai *entropy* terlebih dahulu, dengan menggunakan rumus berikut (Putra et al., 2021) :

Menghitung Entropy

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

Menghitung Nilai Gain

$$\text{Gain}(S, A) = \text{Entropy}(S) - \sum_{i=1}^n \frac{|s_i|}{|S|} * \text{Entropy}(s_i) \quad (2)$$

Keterangan:

S : Total Kasus

n : Banyak partisi S

p_i : Proporsi dari s_i terhadap S

A : Atribut

n : Banyak partisi atribut A

$|s_i|$: Banyak kasus pada partisi ke i

$|S|$: Banyak kasus dalam S

C. Confusion Matrix

Confusion Matrix digunakan untuk mengevaluasi tingkat akurasi hasil klasifikasi secara manual melalui representasi dalam bentuk tabel matriks (Prasojo & Haryatmi, 2021). Confusion matrix menghasilkan tiga metrik evaluasi utama, yaitu *accuracy*, *precision*, dan *recall* (Wati & Rudianto, 2023). Metrik evaluasi utama sebagai berikut (Desyanita & Wibowo, 2020):

1. *Accuracy*, merupakan rasio antara jumlah prediksi yang benar (baik positif maupun negatif) terhadap total keseluruhan data, dihitung sebagai berikut:

$$\frac{TP + TN}{TP + FP + TN + FN} \quad (3)$$

2. *Precision*, merupakan perbandingan antara jumlah prediksi positif yang benar terhadap total data yang diprediksi sebagai positif, dihitung sebagai berikut:

$$\frac{TP}{TP + FP} \quad (4)$$

3. *Recall*, merupakan rasio antara prediksi benar positive dibandingkan dengan keseluruhan data yang benar positive, dihitung sebagai berikut:

$$\frac{TP}{TP + FN} \quad (5)$$

Keterangan:

TP : Jumlah data yang positif dan diprediksi positif.

FP : Jumlah data yang negatif tetapi diprediksi positif.

TN : Jumlah data yang negatif dan diprediksi negatif.

FN : Jumlah data yang positif tetapi diprediksi negatif.

D. Rapid Miner

Rapid Miner adalah perangkat lunak *open source* dengan lisensi AGPL yang berfungsi sebagai alat pengolahan data mining (Puspita et al., 2022).

METODE

Penelitian ini menggunakan pendekatan kuantitatif deskriptif dengan metode klasifikasi data mining algoritma C4.5. Tujuan dari penelitian ini adalah membangun model klasifikasi untuk memprediksi kelayakan pemberian kredit ulang kepada nasabah KSU Bona Mandiri Jaya. Model dibangun berdasarkan data primer yang diperoleh langsung dari koperasi serta diperkuat dengan tinjauan teoritis dari sejumlah artikel ilmiah terkait.

Dalam penelitian ini diperoleh dua jenis data :

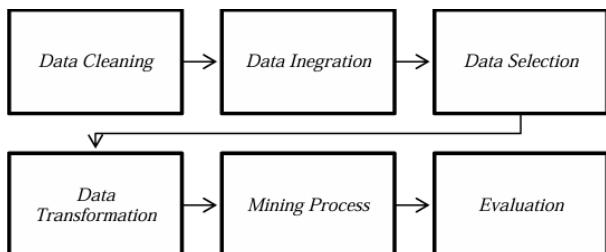
1. Data primer, diperoleh secara langsung dari KSU Bona Mandiri Jaya, Jl. Wan Dahlan Ibrahim (Jl. Merdeka) No. 9-10 Dumai-Riau yang dilaksanakan pada tanggal 6 Februari hingga 15 Maret 2025. Data ini merupakan data historis nasabah peminjam tahun 2022 di wilayah Ressort 1 yang masih aktif sesuai prosedur, yang mencakup informasi

- pinjaman, penghasilan, jaminan, dan status kelayakan.
2. Data sekunder, diperoleh melalui studi literatur dari berbagai artikel ilmiah nasional terkait algoritma C4.5, data mining, dan analisis kelayakan kredit untuk memperkuat landasan teori dan metodologi.

Banyak data pada penelitian ini yaitu 114 data nasabah, terdiri dari beberapa atribut dalam model klasifikasi penilaian kelayakan kredit, yaitu:

1. Riwayat Pinjaman (Lancar, Tidak Lancar, Macet)
2. Jaminan (Tinggi, Sedang, Rendah)
3. Income (Nilai numerik penghasilan perbulan)
4. Status Kelayakan (Layak atau Tidak layak)

Dalam menerapkan pendekatan *data mining* untuk memprediksi potensi risiko kredit, dibutuhkan perencanaan sistematis yang terdiri atas sejumlah tahapan yang dirancang secara terstruktur dan sesuai dengan kebutuhan analisis oleh (Handayani et al., 2021). Proses ini ditampilkan pada Gambar 1 yang mencakup enam langkah utama :



Gambar 1. Tahapan Penelitian

HASIL DAN PEMBAHASAN

A. Perhitungan dengan *Decision Tree* dan Algoritma C4.5

Tabel 1 Sampel Nasabah KSU Bona Mandiri Jaya

No	Nama	Riwayat Pinjaman	Jaminan	Income	Status Kelayakan
1	Sabban Lubis	Tidak lancar	Sedang	6.000.000	Tidak layak
2	Sri Suparmi	Lancar	Rendah	4.000.000	Layak
3	Melinda Puspitasari	Lancar	Sedang	-	Layak
4	Haris Suryadi	Lancar	Sedang	4.000.000	Layak
5	Raja Afrizal	Lancar	Tinggi	5.000.000	Layak
6	Warsito	Macet	Sedang	4.000.000	Tidak layak
7	Amrizal Anara	Lancar	Tinggi	4.000.000	Layak
8	Indra Hifni	Lancar	Tinggi	5.000.000	Layak
9	Punjul Wahyudi	Lancar	Sedang	7.000.000	Layak

10	Ade Senada Harahap	Lancar	Sedang	5.000.000	Layak
11	Upik	Lancar	Rendah	3.000.000	Tidak layak
12	Supriadi	Lancar	Rendah	3.000.000	Layak
13	Nofrizal Akmal	Lancar	Tinggi	6.000.000	Layak
14	Mijar Sihotang	Lancar	Rendah	3.000.000	Layak
15	Ngatino	Lancar	Rendah	2.000.000	Layak
.					
114	Rahmat	Lancar	Tinggi	3.000.000	Layak

Menghitung Entropy dan Nilai Gain

Berdasarkan persamaan (1) dan (2) diperoleh nilai *entropy* dan nilai gain dari setiap atribut sebagai berikut:

Entropy Total:

$$\text{Entropy } (S) =$$

$$\left(\left(-\frac{97}{114} \right) * \log_2 \left(\frac{97}{114} \right) + \left(-\frac{17}{114} \right) * \log_2 \left(\frac{17}{114} \right) \right)$$

$$\text{Entropy } (S) = 0,60764$$

Entropy Atribut:

Iterasi 1

a. Riwayat Pinjaman

$$\text{Entropy } (\text{Lancar}) =$$

$$\left(\left(-\frac{97}{100} \right) * \log_2 \left(\frac{97}{100} \right) + \left(-\frac{3}{100} \right) * \log_2 \left(\frac{3}{100} \right) \right)$$

$$\text{Entropy } (\text{Lancar}) = 0,19439$$

$$\text{Entropy } (\text{Tidak lancar}) =$$

$$\left(\left(-\frac{0}{7} \right) * \log_2 \left(\frac{0}{7} \right) + \left(-\frac{7}{7} \right) * \log_2 \left(\frac{7}{7} \right) \right)$$

$$\text{Entropy } (\text{Tidak lancar}) = 0$$

$$\text{Entropy } (\text{Macet}) =$$

$$\left(\left(-\frac{0}{7} \right) * \log_2 \left(\frac{0}{7} \right) + \left(-\frac{7}{7} \right) * \log_2 \left(\frac{7}{7} \right) \right)$$

$$\text{Entropy } (\text{Macet}) = 0$$

$$\text{Gain}(S, \text{Riwayat pinjaman}) =$$

$$0,60764 - \left(\left(\frac{100}{114} \right) * 0,19439 \right) + \left(\left(\frac{7}{114} \right) * 0 \right) + \left(\left(\frac{7}{114} \right) * 0 \right)$$

$$\text{Gain } (S, \text{Riwayat pinjaman}) = 0,43712$$

b. Jaminan

$$\text{Entropy (Tinggi)} = \left(\left(-\frac{23}{27} \right) * \log_2 \left(\frac{23}{27} \right) + \left(-\frac{4}{27} \right) * \log_2 \left(\frac{4}{27} \right) \right)$$

$$\text{Entropy (Tinggi)} = 0,60519$$

$$\text{Entropy (Sedang)} = \left(\left(-\frac{55}{63} \right) * \log_2 \left(\frac{55}{63} \right) + \left(-\frac{8}{63} \right) * \log_2 \left(\frac{8}{63} \right) \right)$$

$$\text{Entropy (Sedang)} = 0,54911$$

$$\text{Entropy (Rendah)} = \left(\left(-\frac{19}{24} \right) * \log_2 \left(\frac{19}{24} \right) + \left(-\frac{5}{24} \right) * \log_2 \left(\frac{5}{24} \right) \right)$$

$$\text{Entropy (Rendah)} = 0,73828$$

$$\text{Gain}(S, \text{Jaminan}) = 0,60764 - \left(\left(\frac{27}{114} \right) * 0,60519 \right) + \left(\left(\frac{63}{114} \right) * 0,45911 \right) + \left(\left(\frac{24}{114} \right) * 0,73828 \right)$$

$$\text{Gain (S, Jaminan)} = -112,37478$$

c. Income

$$\text{Entropy (> 500.000)} = \left(\left(-\frac{90}{104} \right) * \log_2 \left(\frac{90}{104} \right) + \left(-\frac{14}{104} \right) * \log_2 \left(\frac{14}{104} \right) \right)$$

$$\text{Entropy (> 500.000)} = 0,56996$$

$$\text{Entropy} (\leq 500.000) = \left(\left(-\frac{7}{10} \right) * \log_2 \left(\frac{7}{10} \right) + \left(-\frac{3}{10} \right) * \log_2 \left(\frac{3}{10} \right) \right)$$

$$\text{Entropy} (\leq 500.000) = 0,99129$$

$$\text{Gain (S, Income)} = 0,60764 - \left(\left(\frac{104}{114} \right) * 0,56996 \right) + \left(\left(\frac{10}{114} \right) * 0,99129 \right)$$

$$\text{Gain (S, Income)} = 0,01037$$

Gain terbesar yang diperoleh berada pada atribut Riwayat Keputusan dengan status lancer, maka lanjut ke iterasi 2.

Iterasi 2

a. Jaminan

$$\text{Entropy (Tinggi)} = \left(\left(-\frac{23}{23} \right) * \log_2 \left(\frac{23}{23} \right) + \left(-\frac{0}{23} \right) * \log_2 \left(\frac{0}{23} \right) \right)$$

$$\text{Entropy (Tinggi)} = 0$$

$$\text{Entropy (Sedang)} = \left(\left(-\frac{55}{55} \right) * \log_2 \left(\frac{55}{55} \right) + \left(-\frac{0}{55} \right) * \log_2 \left(\frac{0}{55} \right) \right)$$

$$\text{Entropy (Sedang)} = 0$$

$$\text{Entropy (Rendah)} = \left(\left(-\frac{19}{22} \right) * \log_2 \left(\frac{19}{22} \right) + \left(-\frac{3}{22} \right) * \log_2 \left(\frac{3}{22} \right) \right)$$

$$\text{Entropy (Rendah)} = 0,57464$$

$$\text{Gain (S, Jaminan)} = 0,60764 - \left(\left(\frac{23}{114} \right) * 0 \right) + \left(\left(\frac{55}{114} \right) * 0 \right) + \left(\left(\frac{22}{114} \right) * 0,57464 \right)$$

$$\text{Gain (S, Jaminan)} = -20,19739$$

b. Income

$$\text{Entropy (> 500.000)} = \left(\left(-\frac{90}{90} \right) * \log_2 \left(\frac{90}{90} \right) + \left(-\frac{0}{90} \right) * \log_2 \left(\frac{0}{90} \right) \right)$$

$$\text{Entropy (> 500.000)} = 0$$

$$\text{Entropy} (\leq 500.000) = \left(\left(-\frac{7}{10} \right) * \log_2 \left(\frac{7}{10} \right) + \left(-\frac{3}{10} \right) * \log_2 \left(\frac{3}{10} \right) \right)$$

$$\text{Entropy} (\leq 500.000) = 0,88129$$

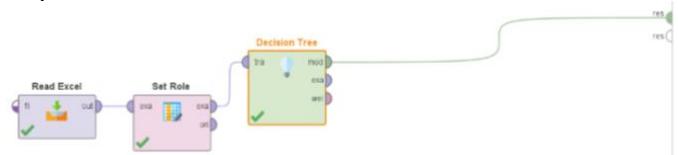
$$\text{Gain (S, Income)} = 0,60764 - \left(\left(\frac{90}{114} \right) * 0,08807 \right) + \left(\left(\frac{10}{114} \right) * 0,72193 \right)$$

$$\text{Gain (S, Income)} = 0,53033$$

Gain terbesar yang diperoleh berada pada atribut Income dengan status ≤ 500.000 , maka atribut terakhir sebagai penentu kelayakan nasabah adalah Jaminan.

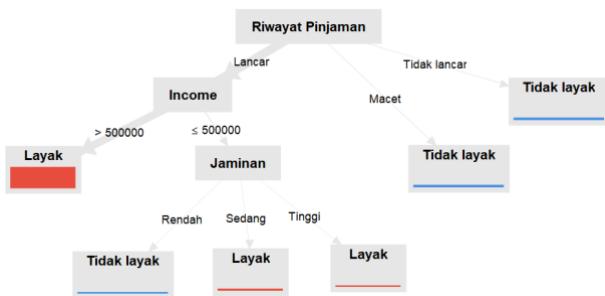
B. Penggunaan Rapid Miner

Dari data yang diterima sebanyak 114 data, peneliti melakukan pengujian data tersebut dengan Rapid Miner.



Gambar 2. Model Decision Tree di Rapid Miner

Berikut hasil *Decision Tree* yang diperoleh dari proses pengolahan data dengan algoritma C4.5 untuk menentukan kelayakan kredit nasabah, yang ditampilkan pada Gambar 3.



Gambar 3. Hasil Model Decision Tree di Rapid Miner

Adapun hasil Deskripsi dari *Decision Tree* diatas yaitu:

```

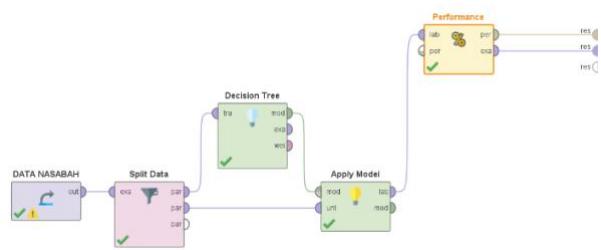
Riwayat Pinjaman = Lancar
| Income > 500000: Layak {Tidak layak=0, Layak=90}
| Income ≤ 500000
| | Jaminan = Rendah: Tidak layak {Tidak layak=3, Layak=0}
| | Jaminan = Sedang: Layak {Tidak layak=0, Layak=5}
| | Jaminan = Tinggi: Layak {Tidak layak=0, Layak=2}
Riwayat Pinjaman = Macet: Tidak layak {Tidak layak=7, Layak=0}
Riwayat Pinjaman = Tidak lancar: Tidak layak {Tidak layak=7, Layak=0}
  
```

Gambar 4. Hasil Deskripsi Decision Tree di Rapid Miner

Berdasarkan *decision tree* yang dihasilkan oleh algoritma C4.5, aturan klasifikasi dapat diturunkan dalam bentuk pernyataan IF-THEN sebagai berikut:

1. IF Riwayat Pinjaman = Lancar AND Income > 500000 THEN Layak
2. IF Riwayat Pinjaman = Lancar AND Income ≤ 500000 AND Jaminan = Rendah THEN Tidak Layak
3. IF Riwayat Pinjaman = Lancar AND Income ≤ 500000 AND Jaminan = Sedang THEN Layak
4. IF Riwayat Pinjaman = Lancar AND Income ≤ 500000 AND Jaminan = Tinggi THEN Layak
5. IF Riwayat Pinjaman = Macet THEN Tidak Layak
6. IF Riwayat Pinjaman = Tidak lancar THEN Tidak Layak

a. Evaluasi dan Validasi Data



Gambar 5. Model Validasi Data *Decision Tree* di Rapid Miner

accuracy: 98.25%			
	true Tidak layak	true Layak	class precision
pred. Tidak layak	16	1	94.12%
pred. Layak	1	96	98.97%
class recall	94.12%	98.97%	

Gambar 6. Hasil Nilai *accuracy*, *precision*, dan *recall* di Rapid Miner

Evaluasi performa model klasifikasi C4.5 dilakukan dengan *confusion matrix* juga beberapa metrik evaluasi seperti akurasi, presisi, dan recall. Berdasarkan hasil evaluasi yang diperoleh, model menghasilkan akurasi sebesar 98.25%, menunjukkan sebagian besar prediksi yang dilakukan oleh model sesuai dengan nilai sebenarnya.

Dari hasil *Confusion Matrix* tersebut dapat dijelaskan bahwa:

1. Berhasil diklasifikasikan dengan benar sebanyak 16 data sebagai tidak layak,
2. Berhasil diklasifikasikan dengan benar sebanyak 96 data sebagai layak,
3. Terdapat 1 data yang seharusnya tidak layak namun diprediksi sebagai layak,
4. Dan terdapat 1 data yang seharusnya layak namun diprediksi sebagai tidak layak .

Menghitung *Accuracy*, dari persamaan (3) kita peroleh :

$$\frac{96 + 16}{96 + 1 + 16 + 1} = \frac{112}{114} = 0,9825$$

Menghitung *Precision* :

$$P(1) = \frac{TP}{TP+FP}, \quad P(0) = \frac{TN}{TN+FN}$$

$$\frac{96}{97} = 0,9897 \quad \frac{16}{17} = 0,9412$$

Menghitung *Recall* :

$$R(1) = \frac{TP}{TP+FN} = \frac{96}{97} = 0,9897$$

$$R(0) = \frac{TN}{TN+FP} = \frac{16}{17} = 0,9412$$

Nilai precision dan recall untuk masing-masing kelas adalah sebagai berikut:

Precision untuk kelas tidak layak	: 94.12%
Precision untuk kelas layak	: 98.97%
Recall untuk kelas tidak layak	: 94.12%
Recall untuk kelas layak	: 98.97%

Adapun rincian hasil *Confusion Matrix Algorithm Decision Tree C4.5* menggunakan data pelatihan sebesar 80% dan data pengujian sebesar 20% mendapatkan nilai akurasi 95,45% dengan kedalaman maksimal tes hasil 10.

PENUTUP

SIMPULAN

Berdasarkan temuan penelitian ini, dapat disimpulkan bahwa penelitian ini berhasil membangun model klasifikasi kelayakan kredit ulang untuk nasabah KSP Bona Mandiri Jaya dengan menggunakan algoritma C4.5 berbasis data mining. Berdasarkan pengujian terhadap 114 data nasabah yang melibatkan atribut riwayat pinjaman, jaminan, dan penghasilan, algoritma C4.5 mampu membentuk pohon keputusan yang menghasilkan aturan klasifikasi yang sederhana dan mudah diinterpretasikan. Model ini kemudian divalidasi menggunakan metode split validation dengan proporsi 80%:20% melalui perangkat lunak Rapid Miner.

Hasil evaluasi menunjukkan model klasifikasi memiliki tingkat akurasi sebesar 98,25%, dengan nilai precision dan recall yang tinggi untuk kedua kelas (layak dan tidak layak), yaitu masing-masing 94,12% untuk kelas "tidak layak" dan 98,97% untuk kelas "layak". Temuan ini menandakan model memiliki akurasi prediksi yang unggul serta konsistensi yang baik dalam membedakan nasabah yang layak dan tidak layak menerima kredit ulang. Model ini diharapkan dapat membantu pihak koperasi dalam menekan risiko gagal bayar serta meningkatkan efisiensi dan objektivitas dalam proses evaluasi kredit nasabah.

SARAN

Penelitian ini masih menggunakan jumlah data yang relatif terbatas, sehingga penelitian berikutnya disarankan untuk menggunakan dataset yang lebih besar, mencakup periode waktu lebih panjang, serta melibatkan lebih banyak atribut agar hasil prediksi lebih akurat dan general. Peneliti selanjutnya juga dapat membandingkan algoritma C4.5 dengan algoritma lain seperti Random Forest atau Naïve Bayes untuk mengetahui metode yang paling baik dalam prediksi kelayakan kredit.

DAFTAR PUSTAKA

- R. D. Perkasa, D. Safitri, I. Lubis, and Z. Aimansyah, "The Role of Savings and Loan Cooperatives in the Development of Micro, Small, and Medium Enterprises (MSMEs)," *Int. J. Educ. Soc. Stud. Manag.*, vol. 4, no. 2, pp. 655–670, 2024, doi: 10.52121/ijessm.v4i2.336.
- "Undang-Undang Nomor 25 Tahun 1992 Tentang Perkoperasian," 1992.
- D. R. W. Maulida, A. P. A. Triyas, and H. L. Putri, "Dinamika Keterlekatkan Ekonomi: Hubungan UMKM dengan Koperasi Simpan Pinjam dalam Meningkatkan Pertumbuhan Usaha," *Arus J. Sos. dan Hum.*, vol. 4, no. 1, pp. 322–329, 2024, doi: 10.57250/ajsh.v4i1.404.
- E. Endris, "Loan repayment performance of micro and small-scale enterprise: evidence from North Wollo Zone, Ethiopia," *Heliyon*, vol. 8, no. 12, p. e12085, 2022, doi: 10.1016/j.heliyon.2022.e12085.
- C. Nisak, R. Firdaus, dan Sampir Andrean Sukoco, S. Andrean Sukoco, and S. Tinggi Ilmu Administrasi Pembangunan Jember, "Analisis Sistem Keputusan Pemberian Kredit Bagi Nasabah Pada Koperasi Simpan Pinjam (KSP) Mulia Jaya," vol. 2, no. 6, pp. 1804–1823, 2024, [Online]. Available: <https://journal.institutkom-edu.org/index.php/multiple>
- C. Christnatalis, R. R. Saragih, and B. C. Tambunan, "Data Mining Algorithm C4.5 Classification Determination Credit Eligibility for Jaya Bersama Cooperatives (Korjabe)," *JURTEKSI (Jurnal Teknol. dan Sist. Informasi)*, vol. 8, no. 1, pp. 59–68, 2021, doi: 10.33330/jurteksi.v8i1.1298.
- D. N. Sholihaningtias, "Rekomendasi Kelayakan

- Penerima Kredit Menggunakan Metode TOPSIS dengan Pembobotan ROC," *J. SAINTEKOM*, vol. 13, no. 1, pp. 88-99, 2023, doi: 10.33020/saintekom.v13i1.376.
- Y. Oktafriani, G. Firmansyah, B. Tjahjono, and A. M. Widodo, "Analysis of Data Mining Applications for Determining Credit Eligibility Using Classification Algorithms C4.5, Naïve Bayes, K-NN, and Random Forest," *Asian J. Soc. Humanit.*, vol. 1, no. 12, pp. 1139-1158, 2023, doi: 10.59888/ajosh.v1i12.119.
- D. Puspita, S. Aminah, and A. Arif, "Prediction System for Credit Eligibility Using C4.5 Algorithm," *J. Informatics Telecommun. Eng.*, vol. 6, no. 1, pp. 148-156, 2022, doi: 10.31289/jite.v6i1.7311.
- R. Sutjiadi, T. Rahmawati, and A. A. Prahartiwi, "Algoritma C4.5 Untuk Klasifikasi Kelayakan Kredit Calon Debitur Pada Sistem Informasi Penjualan," *J. Sist. Inf. Bisnis*, vol. 12, no. 2, pp. 89-97, 2022, doi: 10.21456/vol12iss2pp89-97.
- K. Tsalatsatun Nur Rohmah, "Penerapan Data Mining Algoritma C4.5 Dalam Mengidentifikasi Risiko Kredit Pada Lembaga Keuangan," vol. 2, no. 1, pp. 9-16, 2024.
- M. R. Wirasena and J. Warmansyah, "Penerapan Algoritma C4.5 Untuk Prediksi Kelayakan Pengajuan Kartu Kredit Visa Bagi Nasabah," *Teknols J. Ilm. Teknol. Inf. dan Sains*, vol. 14, no. 2, pp. 296-302, 2024, doi: 10.36350/jbs.v14i2.266.
- A. Gunawan and M. S. Hasibuan, "Perbandingan Algoritma C4.5 dan Naive Bayes Untuk Menentukan Penerima Bantuan Bedah Rumah di Kabupaten Tanggamus," *J. SIMADA (Sistem Inf. dan Manaj. Basis Data)*, vol. 6, no. 1, pp. 12-23, 2023, doi: 10.30873/simada.v6i1.3430.
- M. Martin and L. Nilawati, "Comparison of C4.5 and Naïve Bayes Algorithms for Assessment of Public Complaints Services," *J. Informatics Telecommun. Eng.*, vol. 5, no. 1, pp. 101-111, 2021, doi: 10.31289/jite.v5i1.5292.
- B. de Ville, *Introduction to Data Mining*. 2001. doi: 10.1016/b978-155558242-5/50003-6.
- R. Muttaqien, M. G. Pradana, and A. Pramuntadi, "Implementation of Data Mining Using C4.5 Algorithm for Predicting Customer Loyalty of PT. Pegadaian (Persero) Pati Area Office," *Int. J. Comput. Inf. Syst.*, vol. 2, no. 3, pp. 64-68, 2021, doi: 10.29040/ijcis.v2i3.36.
- M. Firmansyah and R. Aufany, "Implementasi Metode Decision Tree Dan Algoritma C4.5 Untuk Klasifikasi Data Nasabah Bank," *Infokam*, vol. XII, no. 1, pp. 1-12, 2016.
- A. Waluyo, H. Jatnika, M. R. S. Permatasari, T. Tuslaela, I. Purnamasari, and A. P. Windarto, "Data Mining Optimization uses C4.5 Classification and Particle Swarm Optimization (PSO) in the location selection of Student Boardinghouses," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 874, no. 1, 2020, doi: 10.1088/1757-899X/874/1/012024.
- R. S. Putra, E. D. Putra, M. H. Rifqo, and H. Witriyono, "Klasifikasi Penyebaran Covid-19 Menggunakan Algoritma C4.5 Kota Pagar Alam," *JUKOMIKA (Jurnal Ilmu Komput. dan Inform.)*, vol. 4, no. 1, pp. 23-35, 2021, doi: 10.54650/jukomika.v4i1.346.
- B. Prasodo and E. Haryatmi, "Analisa Prediksi Kelayakan Pemberian Kredit Pinjaman dengan Metode Random Forest," *J. Nas. Teknol. dan Sist. Inf.*, vol. 7, no. 2, pp. 79-89, 2021, doi: 10.25077/teknosi.v7i2.2021.79-89.
- E. F. Wati and B. Rudianto, "Penerapan Algoritma KNN, Naive Bayes Dan C4.5 Dalam Memprediksi Kelulusan Mahasiswa," *Format J. Ilm. Tek. Inform.*, vol. 11, no. 2, p. 168, 2023, doi: 10.22441/format.2022.v11.i2.009.
- L. Desyanita and A. Wibowo, "Pemodelan Sistem Prediksi Kelayakan Pengajuan Kredit," *Elkom Elektron. Dan Komput.*, vol. 13, no. 2, pp. 10-22, 2020, [Online]. Available: <https://journal.stekom.ac.id/index.php/elkom/article/view/212>
- A. D. Putri, F. Sholekhah, and E. Dadynata, "The Application of C4.5 Decision Tree Algorithm for Predicting the Survival Rate of Thyroid Cancer Patients Penerapan Algoritma Decesion Tree C4.5 untuk Memprediksi Tingkat Kelangsungan Hidup Pasien Kanker Tiroid," vol. 4, no. October, pp. 1485-1495, 2024.
- Fatmawati, "Perbandingan Algoritma Klasifikasi Data Mining Model C4.5 Dan Naive Bayes Untuk Prediksi Penyakit Diabetes," *J. Techno Nusa Mandiri*, vol. XIII, no. 1, p. 50, 2016.
- N. Handayani, H. Wahyono, J. Trianto, and D. S.

Permana, "Prediksi Tingkat Risiko Kredit dengan Data Mining Menggunakan Algoritma Decision Tree C.45," *JURIKOM (Jurnal Ris. Komputer)*, vol. 8, no. 6, p. 198, 2021, doi: 10.30865/jurikom.v8i6.3643.