

Model Prediksi Risiko Kredit Nasabah Menggunakan Algoritma C4.5

Zaehol Fatah¹, Uny Khafifah²

^{1,2} Universitas Ibrahimy, Situbondo

Email : ¹zaeholfatah@gmail.com , ^{*2}unykhafifah27@gmail.com

Abstrak

Sektor keuangan memiliki peran penting dalam memberikan fasilitas kredit kepada nasabah, baik individu maupun perusahaan. Namun, setiap keputusan pemberian kredit mengandung tingkat risiko tertentu, terutama terkait kemungkinan nasabah gagal membayar kewajibannya. Oleh karena itu, diperlukan metode yang efektif untuk mengelola dan memprediksi risiko kredit. Salah satu metode yang banyak digunakan adalah algoritma Decision Tree C4.5, yang mampu menghasilkan model pengambilan keputusan yang mudah dipahami dan diinterpretasikan. Penelitian ini menerapkan algoritma C4.5 untuk membangun model prediksi risiko kredit nasabah, dengan tujuan membantu lembaga keuangan dalam mengevaluasi kelayakan kredit secara lebih akurat dan objektif. Proses klasifikasi dilakukan dengan mengelompokkan data nasabah ke dalam kategori risiko tertentu berdasarkan atribut-atribut seperti pendapatan, lama bekerja, jumlah pinjaman, dan status pembayaran sebelumnya. Hasil eksperimen menunjukkan bahwa algoritma Decision Tree C4.5 mampu membentuk struktur pohon keputusan yang menggambarkan hubungan antar atribut dan menghasilkan aturan klasifikasi (rule set) yang mudah diinterpretasikan. Dari pengujian yang dilakukan, model memperoleh tingkat akurasi sebesar 82,75%, menunjukkan bahwa metode ini cukup efektif dalam mengklasifikasikan tingkat risiko kredit nasabah.

Kata Kunci: Prediksi Risiko Kredit, Algoritma C4.5, Decision Tree, Data Mining, RapidMiner.

PENDAHULUAN

Seiring dengan pesatnya perkembangan dan pertumbuhan bisnis, persoalan perkreditan sangat menarik untuk dibahas dan dipelajari (Ariadi 2020). Kredit adalah penyediaan uang atau tagihan yang dapat dipersamakan dengan itu, berdasarkan persetujuan pinjam meminjam antara bank dengan pihak lain yang mewajibkan pihak peminjam melunasi hutangnya setelah jangka waktu tertentu dengan pemberian bunga (Eka Yuni Titik Artaningsih and Abdullah Muhajir 2023).

Kredit sekarang menjadi tren di masyarakat. Problem kredit adalah sejarah penggunaan kartu kredit yang salah. Dampak yang ditimbulkan dapat menyebabkan kredit macet. Jika pelanggan tidak membayar utang yang telah disepakati dengan bank, dapat meningkatkan risiko kredit yang ada (Pahlevi, Rianto, and Author 2025). Proses

persetujuan kartu kredit merupakan tantangan signifikan bagi institusi keuangan karena melibatkan analisis berbagai faktor, seperti pendapatan, riwayat kredit, dan komitmen finansial calon nasabah (Puspa et al. 2025). Untuk mengurangi risiko kredit, lembaga keuangan melakukan analisis data sebelum memberikan kredit. Nasabah dapat merugikan perusahaan apabila tidak membayar kredit dengan lancar (Kredit and Lembaga 2024).

Nasabah adalah pelanggan (customer) yaitu individu atau perusahaan yang mendapatkan manfaat atau produk dan jasa dari sebuah (lembaga keuangan). Menurut Pardede (2004:9), nasabah adalah orang yang mempercayakan pengurusan uangnya kepada bank untuk digunakan dalam operasional bisnis perbankan. nasabah yang terpilih dalam pengajuan kartu kredit masih kurang tepat dan

penyeleksian nasabah kurang maksimal sehingga nasabah yang telah terpilih tersebut tidak sesuai dengan kriteria yang sudah ditetapkan. Berdasarkan pada permasalahan yaitu tidak akurat dan tidak efektifnya prediksi kelayakan pengajuan kartu kredit bagi nasabah (Wirasena and Warmansyah 2024). Oleh karena itu untuk mengatasi permasalahan itu pihak perbankan atau lembaga keuangan perlu memprediksi kelayakan pengajuan pembiayaan terlebih dahulu. Teknik data mining digunakan untuk menentukan prediksi kelayakan pengajuan pembiayaan dalam membuat keputusan kepada nasabah yang berhak diterima atau ditolak pengajuannya (Kusmira 2021).

Salah satu metode yang banyak digunakan dalam data mining untuk tujuan klasifikasi dan prediksi adalah algoritma C4.5. Algoritma ini merupakan pengembangan dari metode pohon keputusan (*Decision Tree*) yang bekerja dengan menghitung nilai information gain untuk menentukan atribut terbaik dalam membentuk model klasifikasi (Nasrullah 2021). Setiap cabang pohon keputusan menggambarkan suatu kondisi atau atribut tertentu, sedangkan daun pohon menunjukkan hasil atau kategori keputusan. Keunggulan algoritma C4.5 dibanding metode lain adalah kemampuannya dalam menangani atribut data yang bersifat kontinu maupun diskrit, mengatasi data yang hilang (*missing value*), serta menghasilkan aturan keputusan (*rule set*) yang mudah dipahami dan diinterpretasikan oleh pengambil keputusan non teknis (Afriansyah and Annisa 2022). Dengan demikian, algoritma ini sangat cocok diterapkan pada permasalahan klasifikasi risiko kredit yang melibatkan banyak variabel nasabah.

Dalam penelitian ini, algoritma C4.5 diterapkan untuk membangun model prediksi risiko kredit nasabah dengan menggunakan dataset yang memuat berbagai atribut penting seperti usia, pendapatan, status pekerjaan, jumlah pinjaman, riwayat pembayaran, status

perkawinan, serta durasi pinjaman. Hasil dari penelitian diharapkan tercipta model prediksi risiko kredit yang efektif, transparan, dan mudah diterapkan oleh lembaga keuangan. Model ini tidak hanya bermanfaat bagi pihak bank dalam menilai kelayakan nasabah, tetapi juga dapat digunakan sebagai alat bantu pengawasan dalam manajemen risiko kredit secara keseluruhan. Pada akhirnya, penerapan algoritma C4.5 dalam analisis risiko kredit diharapkan mampu meningkatkan efisiensi operasional lembaga keuangan, memperkuat stabilitas sistem keuangan, serta memberikan kontribusi nyata terhadap pembangunan ekonomi yang berkelanjutan.

METODE

Dalam dunia akademik, "penelitian" didefinisikan sebagai pekerjaan yang rajin dan atau penyelidikan tentang suatu topik untuk menemukan atau merevisi fakta, teori, aplikasi, dengan tujuan untuk menemukan dan menyebarkan pengetahuan baru. Metode penelitian yang dilakukan adalah metode penelitian eksperimen, dengan tahapan sebagai berikut:

1. Pengumpulan Data

Pada tahap ini, data yang akan diproses dipilih; kemudian, mencari data tambahan yang diperlukan; dan akhirnya, menggabungkan semua data ke dalam set data, termasuk variabel yang diperlukan untuk proses.

2. Pengolahan Data Awal (Data Pre-processing)

Pada tahap ini, data dipilih, dibersihkan, dan diubah ke bentuk yang diinginkan untuk memungkinkan persiapan untuk pembuatan model.

3. Model

Pada tahap ini, data dianalisis dan dikelompokkan berdasarkan hubungannya. Setelah data dianalisis, model-model yang sesuai diterapkan. Untuk membuat model, juga diperlukan pembagian data ke dalam data latihan

(training data) dan data uji (Testing Data).

4. Eksperimen dan Pengujian Metode (Method Test and Experiment)

Pada tahap ini dilakukan pengujian terhadap model yang diusulkan untuk mengetahui sejauh mana tingkat keakuratan model dalam melakukan klasifikasi dan menghasilkan aturan keputusan yang tepat. Pengujian dilakukan menggunakan operator Split Data untuk membagi dataset menjadi data latih (training) dan data uji (testing), serta operator Apply Model dan Performance (Classification) untuk mengevaluasi hasil prediksi yang dihasilkan oleh algoritma C4.5 (Novianti, Mandati, and Andana 2023).

Pengumpulan data

Data yang digunakan dalam penelitian ini adalah Credit Risk Dataset, yang diperoleh dari situs Kaggle (<https://www.kaggle.com/datasets/laotse/credit-risk-dataset>)(Anon n.d.). Dataset ini merupakan data terbuka (open dataset) yang sering digunakan dalam penelitian terkait prediksi risiko kredit nasabah menggunakan metode klasifikasi. Dataset ini berisi informasi mengenai karakteristik pribadi, kondisi keuangan, serta histori kredit nasabah yang digunakan untuk memprediksi apakah seorang nasabah berpotensi mengalami gagal bayar (default) atau tidak (Saputra, Atina, and Nastiti 2024). Dataset ini terdiri dari 32.581 data nasabah (record) dengan 10 atribut (variabel) yang merepresentasikan faktor-faktor penting dalam penilaian risiko kredit.

Atribut-atribut yang terdapat pada dataset ini meliputi:

1. person_age: Usia nasabah.
2. person_income: Pendapatan tahunan nasabah.
3. person_home_ownership: Status kepemilikan rumah.
4. person_emp_length: Lama bekerja.
5. loan_intent: Tujuan pengajuan pinjaman.
6. loan_grade: Peringkat kualitas pinjaman berdasarkan profil risiko.
7. loan_amnt: Jumlah pinjaman yang diajukan.
8. loan_int_rate: Suku bunga pinjaman.
9. cb_person_default_on_file: Riwayat gagal bayar.
10. cb_person_cred_hist_length: Lama riwayat kredit.

Pengolahan Data Awal

Data yang didapatkan dari sumber merupakan data asli dari situs Kaggle yang belum mengalami proses perubahan. Dataset tersebut kemudian dilakukan proses data cleaning atau perapihan agar lebih mudah dibaca dan diproses dalam tahap analisis. Proses perapihan ini mencakup penghapusan data kosong (missing values), penyesuaian tipe data (numerik dan kategorikal), serta normalisasi beberapa atribut numerik agar hasil prediksi menjadi lebih akurat (Faurika, Naseh Khudori, and Haris 2024). Dataset Kaggle ini memudahkan peneliti dalam tahap persiapan data karena sudah memiliki struktur yang jelas dan relevan untuk analisis risiko kredit nasabah.

Tabel 1 Dataset Risiko Kredit Nasabah Pengolahan Data Awal

No	Nama Atribut	Jenis Data	Keterangan
1.	person_age	Numerik	Usia nasabah
2.	person_income	Numerik	Pendapatan tahunan nasabah
3.	person_home_ownership	Kategorikal	Status kepemilikan rumah
4.	person_emp_length	Numerik	Lama bekerja nasabah
5.	loan_intent	Kategorikal	Tujuan pinjaman
6.	loan_grade	Kategorikal	kualitas pinjaman berdasarkan risiko
7.	loan_amnt	Numerik	Jumlah pinjaman yang diajukan (USD)
8.	loan_int_rate	Numerik	Suku bunga pinjaman (%)
9.	cb_person_default_on_file	Kategorikal	Status gagal bayar

No Nama Atribut	Jenis Data	Keterangan
10. cb_person_cred_hist_length	Numerik	Lama riwayat kredit

Metode Penelitian

Metode yang digunakan dalam penelitian ini adalah algoritma C4.5, yang merupakan salah satu teknik klasifikasi berbasis Decision Tree paling populer dalam bidang data mining dan machine learning (Hidayatsyah 2020). Algoritma ini digunakan untuk membangun model klasifikasi yang mampu memprediksi tingkat risiko kredit nasabah berdasarkan atribut-atribut keuangan dan demografis yang tersedia pada dataset. Secara konseptual, Decision Tree bekerja dengan mengubah sekumpulan data (record) menjadi sebuah struktur pohon keputusan (decision tree structure) yang berisi simpul (node), cabang (branch), dan daun (leaf). Setiap simpul merepresentasikan atribut atau variabel prediktor, sedangkan setiap cabang menunjukkan hasil uji kondisi pada atribut tersebut, dan setiap daun menyatakan keputusan atau kelas akhir (misalnya, layak atau tidak layak kredit).

Berikut merupakan rumus entropy untuk mengukur tingkat ketidakpastian himpunan data yang digunakan dalam algoritma c4.5

$$\text{Entropy}(S) = -I = 1 \sum p_i \log_2(p_i)$$

Pengujian Metode

Tahap ini bertujuan untuk menguji tingkat keakuratan model klasifikasi yang telah dibangun menggunakan algoritma C4.5. Pengujian dilakukan menggunakan aplikasi RapidMiner Studio melalui beberapa langkah utama sebagai berikut:

1. Import Data (Read Excel)
Dataset risiko kredit nasabah diimpor menggunakan operator Read Excel. Dataset ini berisi atribut seperti status perkawinan, kepemilikan rumah, pekerjaan, jumlah tanggungan, serta status risiko kredit (target klasifikasi).
2. Menentukan Label (Set Role)
Operator Set Role digunakan untuk menentukan kolom target (label) yang akan diprediksi, yaitu atribut Risiko Kredit. Atribut ini dijadikan label agar dapat dikenali sebagai variabel dependen oleh sistem.
3. Pembagian Data (Split Data)
Dataset dibagi menjadi dua bagian dengan perbandingan 70% data training dan 30% data testing menggunakan operator Split Data. Data training digunakan untuk membentuk model, sedangkan data testing digunakan untuk menguji performa model terhadap data baru.
4. Pembuatan Model (Decision Tree)
Operator Decision Tree digunakan untuk membangun model klasifikasi berdasarkan algoritma C4.5. Proses pemilihan atribut terbaik dilakukan dengan kriteria Gain Ratio, dan pohon keputusan terbentuk berdasarkan nilai entropi tiap atribut.
5. Penerapan Model (Apply Model)
Model yang telah terbentuk dari data training kemudian diterapkan pada data testing menggunakan operator Apply Model. Hasilnya berupa data testing yang sudah diprediksi kelas risikonya oleh sistem.
6. Evaluasi Kinerja Model (Performance Classification).
Dari hasil Performance Vector yang ditampilkan pada RapidMiner, dapat dijelaskan bahwa model klasifikasi yang diuji menunjukkan tingkat kinerja yang baik dalam memprediksi kelas data. Nilai Accuracy menunjukkan persentase data uji yang berhasil diprediksi dengan benar oleh model. Nilai Precision mengindikasikan seberapa akurat

model dalam mengklasifikasikan data positif, sedangkan Recall menggambarkan kemampuan model dalam mengenali seluruh data positif yang sebenarnya.

HASIL DAN PEMBAHASAN

Dalam penelitian ini, metode yang digunakan adalah algoritma C4.5 dengan proses pengolahan data melalui beberapa tahap seleksi dan transformasi data. Data yang digunakan dalam penelitian ini diperoleh secara sekunder dari platform Kaggle yang menyediakan dataset terkait risiko kredit nasabah. Dataset tersebut terdiri dari sejumlah record yang menggambarkan berbagai faktor yang memengaruhi kelayakan kredit, seperti pendapatan, jumlah pinjaman, riwayat pembayaran, durasi pinjaman, status pekerjaan, serta status pernikahan. Untuk memastikan keakuratan hasil klasifikasi, dataset dibagi menjadi dua bagian, yaitu 70% digunakan sebagai data training untuk membangun model, dan 30% digunakan sebagai data testing untuk menguji performa model (Puspa et al. 2025).

Proses pengolahan data dilakukan menggunakan RapidMiner, yang meliputi tahap pembersihan data, transformasi atribut, dan seleksi fitur agar data siap diproses oleh algoritma C4.5. Selanjutnya, RapidMiner digunakan untuk membentuk model pohon keputusan (Decision Tree) berdasarkan data training, kemudian dilakukan evaluasi hasil prediksi menggunakan data testing untuk mengukur tingkat akurasi model (Ardiyansyah, Rahayuningsih, and Maulana 2018).

Pengujian Menggunakan Algoritma C4.5

Transformasi data

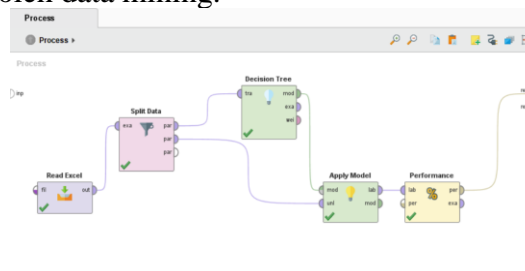
Langkah transformasi data merupakan salah satu tahap penting dalam proses data preprocessing yang bertujuan untuk menyesuaikan format dan struktur data agar siap digunakan pada tahap analisis berikutnya, khususnya dalam penerapan teknik data mining. Transformasi dilakukan agar data yang diperoleh dari berbagai sumber memiliki format yang konsisten, bersih, dan sesuai

dengan kebutuhan algoritma yang akan digunakan, sehingga hasil analisis dapat lebih akurat dan reliabel.

	_SEJA...	HARI_BEKE...	ADA_MOBIL...	ADA_TELE...	ADA_EMAIL...	JENIS_PEK...	JUMLAH_A...
	nominal	polynomial	polynomial	polynomial	polynomial	polynomial	label
1	35	-4542	1	1	0	?	2
2	35	-4542	1	1	0	?	2
3	74	-1134	1	0	0	0	2
4	10	-3051	1	0	1	1	1
5	10	-3051	1	0	1	1	1
6	10	-3051	1	0	1	1	1
7	10	-3051	1	0	1	1	1
8	34	365243	1	0	0	?	1
9	34	365243	1	0	0	?	1
10	34	365243	1	0	0	?	1
11	72	-769	1	1	1	1	2
12	72	-769	1	1	1	1	2
13	72	-769	1	1	1	1	2
14	78	-1194	1	0	0	0	2
15	78	-1194	1	0	0	0	2
16	78	-1194	1	0	0	0	2
17	78	-1194	1	0	0	0	2
18	78	-1194	1	0	0	0	2

Gambar 1. Format Data Proses Pengolahan Data

Pada proses implementasi menggunakan RapidMiner, dataset yang telah melalui tahap seleksi dan transformasi akan dikelola untuk membangun model prediksi risiko kredit nasabah dengan algoritma C4.5. RapidMiner menyediakan berbagai operator yang digunakan untuk proses pengolahan data, pembentukan pohon keputusan, serta evaluasi hasil prediksi guna menentukan tingkat risiko kredit setiap nasabah secara akurat. Pada Gambar. 2 menunjukkan rumus untuk menghitung penggunaan software rapidminer yang didalamnya sudah dilengkapi dengan fitur-fitur yang dibutuhkan oleh data mining.

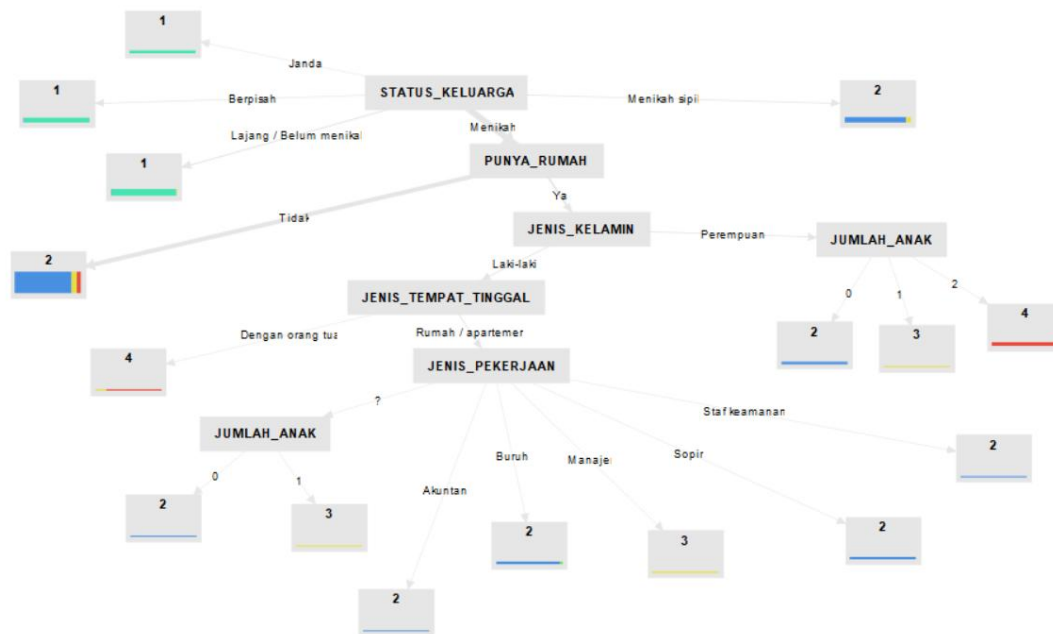


Gambar 2. Pengolahan Data Pohon Keputusan

Hasil dari implementasi algoritma C4.5 menghasilkan pohon keputusan yang merepresentasikan proses klasifikasi risiko kredit nasabah. Pohon keputusan ini menampilkan atribut utama sebagai node

akar (root) yang memiliki pengaruh paling besar terhadap keputusan, diikuti oleh cabang-cabang atribut lain yang menjelaskan kondisi nasabah hingga sampai pada daun (leaf) yang menunjukkan kategori risiko, seperti Risiko Rendah, Risiko Sedang, dan Risiko

Tinggi (Nasrullah 2021). Melalui visualisasi pohon keputusan ini, dapat diketahui pola dan aturan (rule) yang terbentuk dalam proses penentuan tingkat risiko kredit setiap nasabah secara jelas dan terukur.

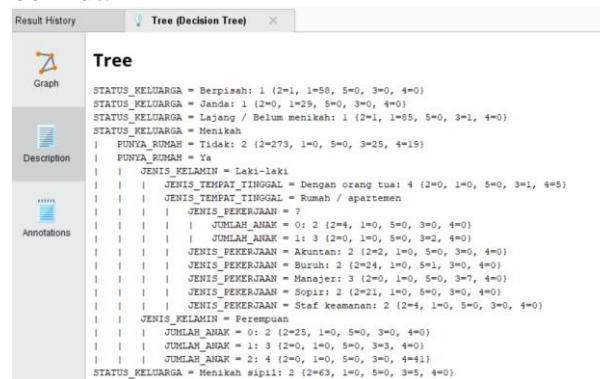


Gambar 3. Pohon Keputusan

Deskripsi Pohon Keputusan

Berdasarkan hasil pohon keputusan (decision tree) yang dihasilkan melalui proses klasifikasi, dapat disimpulkan bahwa terdapat sejumlah aturan (rules) dan kondisi pengambilan keputusan yang terbentuk dari hasil pengolahan dataset. Pohon keputusan ini menggambarkan bagaimana setiap atribut atau variabel dalam dataset berkontribusi terhadap penentuan kelas akhir yang menjadi target analisis. Dengan kata lain, setiap percabangan dalam pohon menunjukkan proses pemilihan atribut dengan nilai gain ratio tertinggi yang memiliki pengaruh paling signifikan terhadap hasil klasifikasi. Hasil pembentukan pohon keputusan ini merupakan keluaran dari proses pengolahan data menggunakan perangkat lunak RapidMiner, yang secara otomatis menganalisis hubungan antar atribut dan

menghasilkan struktur keputusan yang mudah dipahami dalam bentuk visual. Berdasarkan pohon keputusan, dapat disimpulkan bahwa terdapat beberapa kondisi atau aturan yang terbentuk dalam proses klasifikasi. Deskripsi pohon keputusan dari hasil pengolahan data pada RapidMiner dapat dilihat pada Gambar 4 berikut.



Gambar 4. Deskripsi Pohon Keputusan

Hasil gambar di atas dapat dijelaskan bahwa terdapat beberapa atribut yang memiliki penilaian terhadap prediksi risiko kredit nasabah. Dalam hal ini terdapat satu atribut yang paling berpengaruh terhadap prediksi risiko kredit nasabah, yaitu atribut riwayat pembayaran. Hal ini menunjukkan bahwa semakin baik riwayat pembayaran nasabah, maka semakin kecil kemungkinan nasabah tersebut mengalami risiko kredit macet.

Performance Vektor

Berdasarkan hasil Performance Vector (Performance) pada gambar di atas, diperoleh nilai akurasi sebesar 90,33%, yang menunjukkan bahwa model klasifikasi mampu memprediksi data dengan tingkat ketepatan yang cukup tinggi. Dari Confusion Matrix, terlihat bahwa sebagian besar data pada kelas 2 dan kelas 1 berhasil diklasifikasikan dengan benar, dengan nilai recall masing-masing sebesar 98,32% dan 100%. Artinya, hampir seluruh data aktual dari kedua kelas tersebut berhasil dikenali dengan benar oleh model. Selain itu, nilai precision tertinggi juga terlihat pada kelas 1, 3, dan 4 yang mencapai 94,81%, 100%, dan 100%, yang menunjukkan bahwa model memiliki kemampuan yang sangat baik dalam mengklasifikasi data positif pada kelas tersebut. Performance vektor dapat dilihat pada Gambar 5 berikut.

accuracy 90.33%						
	True 2	True 1	True 5	True 3	True 4	class precision
pred 2	178	0	1	11	13	87.50%
pred 1	3	73	0	1	0	94.81%
pred 5	0	0	0	0	0	0.00%
pred 3	0	0	0	7	0	100.00%
pred 4	0	0	0	0	15	100.00%
class recall	98.32%	100.00%	0.00%	38.84%	53.57%	

Gambar 5. Performance Vektor

SIMPULAN

Berdasarkan hasil penelitian yang dilakukan, dapat disimpulkan bahwa penerapan algoritma C4.5 pada model prediksi risiko kredit nasabah mampu menghasilkan klasifikasi yang akurat dan andal dalam menentukan tingkat risiko kredit. Berdasarkan hasil pengujian menggunakan operator Performance

(Classification) di RapidMiner, diperoleh nilai akurasi sebesar 90,33%, yang menunjukkan bahwa model mampu memprediksi tingkat risiko kredit dengan ketepatan yang tinggi. Dari hasil pohon keputusan yang terbentuk, diketahui bahwa atribut seperti riwayat pembayaran, status pernikahan, kepemilikan rumah, jenis pekerjaan, jumlah anak, dan jenis kelamin merupakan faktor-faktor utama yang memengaruhi tingkat risiko kredit nasabah. Nasabah dengan status menikah, memiliki rumah, dan bekerja pada sektor formal seperti akuntan atau staf administrasi cenderung memiliki risiko kredit rendah. Sebaliknya, nasabah dengan status berpisah, tidak memiliki rumah, atau bekerja sebagai buruh dan sopir lebih berpotensi memiliki risiko kredit tinggi. Secara keseluruhan, model Decision Tree dengan algoritma C4.5 terbukti efektif dalam mengidentifikasi pola dan faktor utama yang berpengaruh terhadap risiko kredit nasabah. Hasil ini menunjukkan bahwa metode C4.5 dapat menjadi alat bantu yang objektif, efisien, dan tepat sasaran bagi lembaga keuangan dalam pengambilan keputusan terkait pemberian kredit serta dalam upaya meminimalkan potensi kredit bermasalah (non-performing loan).

DAFTAR PUSTAKA

Afriansyah, Aidil, and Resty Annisa. 2022. "Metode Decision Tree Untuk Memprediksi Kelancaran Angsuran Nasabah Pembiayaan Mikro KCP BSI." *Jurnal Ilmiah Binary STMIK Bina Nusantara Jaya Lubuklinggau* 4(2):48–56. doi: 10.52303/jb.v4i2.75.

Anon. n.d. "Credit Risk Dataset." Retrieved October 17, 2025 (<https://www.kaggle.com/datasets/laotse/credit-risk-dataset>).

Ardiyansyah, Panny Agustia Rahayuningsih, and Reza Maulana. 2018. "Analisis Perbandingan Algoritma Klasifikasi Data Mining Untuk Dataset Blogger Dengan Rapid Miner." *Jurnal Khatulistiwa*

- Informatika* VI(1):20–28.
- Ariadi, Fadly. 2020. “Analisa Perbandingan Algoritma DT C.45 Dan Naïve Bayes Dalam Prediksi Penerimaan Kredit Motor.” *KERNEL: Jurnal Riset Inovasi Bidang Informatika Dan Pendidikan Informatika* 1(1):1–8. doi: 10.31284/j.kernel.2020.v1i1.1183.
- Eka Yuni Titik Artaningsih, and Abdullah Muhajir. 2023. “Komparasi Algoritma Klasifikasi C4.5 Dan C4.5 Berbasis Particle Swarm Optimization Untuk Evaluasi Penentuan Kelayakan Pemberian Kredit Koperasi Di PT. Indah Kiat Pulp & Paper TBK.” *Jurnal Sains, Teknologi Dan Masyarakat* 3(4):29–48.
- Faurika, Faurika, Ahsanun Nash Khudori, and M. Syauqi Haris. 2024. “Application of Cross-Validation Techniques to Handle Overfitting in a Case Study of Decision Tree Implementation for Lung Cancer Prediction.” *ROUTERS: Jurnal Sistem Dan Teknologi Informasi* 2(2):111–20. doi: 10.25181/rt.v2i2.3631.
- Hidayatsyah, Muhammad Rido. 2020. “Penerapan Metode Decision Tree Dalam Pemberian Pinjaman Kepada Debitur.” *PENERAPAN METODE DECISION TREE DALAM PEMBERIAN PINJAMAN KEPADA DEBITUR DENGAN ALGORITMAC4.5 (Studi* 5:22.
- Kredit, Risiko, and Pada Lembaga. 2024. “Tsalatsatun Nur Rohmah, Kurniawati PENERAPAN DATA MINING ALGORITMA C4.5 DALAM MENGIDENTIFIKASI RISIKO KREDIT PADA LEMBAGA KEUANGAN 1,2.” 2(1):9–16.
- Kusmira, Mira. 2021. “Penerapan Data Mining Pengajuan Pembiayaan Perumahan (Consumen Financing) Individual Menggunakan Algoritma C4.5.” *Jurnal Khatulistiwa Informatika* 7(1):66–71. doi: 10.31294/jki.v7i1.126.
- Nasrullah, Asmaul Husnah. 2021. “Implementasi Algoritma Decision Tree Untuk Klasifikasi Produk Laris.” *Jurnal Ilmiah Ilmu Komputer* 7(2):45–51. doi: 10.35329/jiik.v7i2.203.
- Novianti, Triuli, Sri Amaliah Mandati, and Erie Kresna Andana. 2023. “Peningkatan Evaluasi Risiko Kredit Menggunakan Decision Tree C 4.5.” *Journal of Manufacturing in Industrial Engineering & Technology* 2(2):1–9. doi: 10.30651/mine-tech.v2i2.21749.
- Pahlevi, Omar, Harsih Rianto, and Corresponding Author. 2025. “Analisa Komparasi Model Data Mining Algoritma C4.5, CHAID, Dan Random Forest Untuk Penilaian Kelayakan Kredit.” *Computer Science (CO-SCIENCE* 5(1):49–57.
- Puspa, Anindita, Ayu Prayogi, Altha Inas Shofyana, and Dewi Putriani. 2025. “Prediksi Credit Card Approval Menggunakan Algoritma Random.” 2(1).
- Saputra, Dwi Bagus, Vihi Atina, and Faulinda Eli Nastiti. 2024. “Penerapan Model Crisp-Dm Pada Prediksi Nasabah Kredit Menggunakan Algoritma Random Forest.” *IDEALIS : InDonEsiA Journal Information System* 7(2):240–47. doi: 10.36080/idealism.v7i2.3244.
- Wirasena, M. Rifqi, and Julio Warmansyah. 2024. “Penerapan Algoritma C4.5 Untuk Prediksi Kelayakan Pengajuan Kartu Kredit Visa Bagi Nasabah.” *TeknoIS : Jurnal Ilmiah Teknologi Informasi Dan Sains* 14(2):296–302. doi: 10.36350/jbs.v14i2.266.