

## Peningkatan Evaluasi Risiko Kredit Menggunakan Decision Tree C 4.5

Triuli Novianti<sup>1</sup>, Sri Amaliah Mandati<sup>2</sup>, Erie Kresna Andana<sup>3</sup>

Department of Electrical Engineering, Universitas Muhammadiyah Surabaya<sup>1</sup>

Department of Industrial Engineering, Universitas Muhammadiyah Surabaya<sup>2</sup>

Department of Informatics, Universitas Muhammadiyah Surabaya<sup>3</sup>

[triuli.novianti@ft.um-surabaya.ac.id](mailto:triuli.novianti@ft.um-surabaya.ac.id)<sup>1</sup> [sri.amaliah.mandati@um-surabaya.ac.id](mailto:sri.amaliah.mandati@um-surabaya.ac.id)<sup>2</sup>

[erie.kresna@um.surabaya.ac.id](mailto:erie.kresna@um.surabaya.ac.id)<sup>3</sup>

---

### Article Information

#### Article history:

Submitted November 21, 2023

Accepted December 14, 2023

Published December 28, 2023

---

#### Keyword:

Sektor Keuangan

Manajemen Risiko Kredit

Decision Tree C 4.5

---

### ABSTRAK

Sektor keuangan sangat penting dalam memberikan kredit kepada individu dan perusahaan, namun keputusan kredit pada dasarnya berisiko. Metode manajemen risiko kredit yang efektif sangatlah penting. Metode Decision Tree C 4.5 menawarkan model pengambilan keputusan yang dapat ditafsirkan, yang penelitian ini terapkan pada data Kredit Jerman untuk meningkatkan sistem evaluasi risiko kredit. Klasifikasi mengelompokkan data ke dalam kelas berdasarkan atribut, menggunakan algoritma pembelajaran mesin untuk memprediksi label kelas. Penambahan data mengungkap pola dalam kumpulan data besar melalui langkah-langkah seperti pengumpulan data, pembersihan, pemilihan fitur, dan pemodelan. Decision Tree mewakili keputusan dan hasil, dengan Decision Tree C 4.5 menjadi algoritma penting untuk tugas klasifikasi. Peneliti tersebut menggunakan Penguatan Informasi untuk pemilihan atribut, membagi data pada node, dan dapat menangani data yang hilang dan pemangkasan. Kredit merupakan alat keuangan yang harus dikelola untuk memitigasi risiko. RapidMiner adalah platform sumber terbuka untuk analisis data, menampilkan antarmuka yang ramah pengguna, pustaka algoritma, dan alat validasi model. Penelitian eksperimental ini mengikuti tahapan pengumpulan data, praproses, usulan model, pengujian, dan evaluasi hasil. Data kredit Jerman dengan 1000 catatan dan 20 atribut digunakan. Algoritma Decision Tree C4.5 diterapkan menggunakan RapidMiner, dengan data dibagi menjadi set pelatihan dan pengujian untuk menentukan akurasi tertinggi. Algoritma Decision Tree C 4.5 yang diterapkan di RapidMiner menghasilkan Decision Tree dan kumpulan aturan. Akurasi terbaik dicapai dengan 75% data latih dan 25% data uji, sehingga menghasilkan akurasi 71,60%. Metode Decision Tree C 4.5 secara efektif mengklasifikasikan data kredit, memberikan wawasan berharga untuk evaluasi risiko kredit. Penelitian ini berkontribusi pada pengembangan model penerimaan kredit yang andal, bermanfaat bagi lembaga keuangan dan masyarakat.

© This work is licensed under a Creative Commons Attribution 4.0 International License.

---

#### \*Corresponding Author:

Triuli Novianti

Department of Electrical Engineering

Universitas Muhammadiyah Surabaya

Jalan Sutorejo 59, Surabaya, Indonesia

Email: [triuli.novianti@ft.um-surabaya.ac.id](mailto:triuli.novianti@ft.um-surabaya.ac.id)

## 1. PENDAHULUAN

Di tengah dinamika ekonomi yang terus berkembang, sektor keuangan memegang peran krusial dalam memberikan dukungan keuangan kepada individu maupun perusahaan. Pemberian kredit menjadi salah satu instrumen utama yang memungkinkan masyarakat untuk mewujudkan berbagai proyek dan kebutuhan finansial. Meskipun demikian, pengambilan keputusan terkait penerimaan kredit tidak terlepas dari risiko. Oleh karena itu, penting untuk mengembangkan metode yang efektif dalam mengidentifikasi dan mengelola risiko kredit.

Salah satu pendekatan yang menjanjikan dalam meningkatkan proses pengambilan keputusan terkait penerimaan kredit adalah dengan memanfaatkan metode Decision Tree C 4.5 dalam analisis data. Keunggulan Decision Tree C 4.5 terletak pada kemampuannya untuk menghasilkan model keputusan yang mudah diinterpretasi, yang memungkinkan para praktisi dan pengambil keputusan untuk memahami faktor-faktor kunci yang memengaruhi hasil klasifikasi.

Penelitian ini difokuskan pada klasifikasi penerimaan kredit dengan menggunakan metode Decision Tree C.45, dengan data yang diperoleh dari open data German Credit. Data ini menyediakan informasi yang kaya dan terstruktur mengenai profil kreditur, termasuk riwayat kredit, status pekerjaan, kondisi keuangan, dan faktor-faktor lain yang memengaruhi keputusan penerimaan kredit.

Melalui pemanfaatan data terbuka German Credit, penelitian ini bertujuan untuk mengembangkan model Decision Tree C 4.5 yang dapat memberikan gambaran yang akurat dan dapat diandalkan terkait kelayakan penerimaan kredit. Hasil dari penelitian ini diharapkan dapat memberikan kontribusi signifikan terhadap pemahaman dan pengembangan sistem evaluasi risiko kredit, memberikan manfaat baik bagi lembaga keuangan maupun masyarakat pada umumnya.

## 2. DASAR TEORI

### 1. Definisi Klasifikasi

Klasifikasi adalah proses pengelompokan atau pengategorian objek atau data ke dalam kelas atau kategori yang sudah ditentukan berdasarkan karakteristik atau atribut tertentu. Dalam konteks machine learning dan data mining, klasifikasi adalah salah satu tugas yang melibatkan pembangunan model yang dapat memprediksi kelas atau label suatu objek berdasarkan fitur atau atribut yang dimilikinya [13].

Proses klasifikasi biasanya melibatkan penggunaan algoritma machine learning atau teknik statistik untuk memahami pola-pola dalam data yang dapat membedakan antara kelas-kelas yang berbeda. Model yang dihasilkan dari proses klasifikasi dapat digunakan untuk mengklasifikasikan data baru yang belum dikenal ke dalam salah satu kelas yang telah ditentukan sebelumnya [9].

Contoh penerapan klasifikasi dalam kehidupan sehari-hari termasuk pengenalan pola wajah, klasifikasi spam dalam email, atau bahkan prediksi penerimaan kredit seperti yang disebutkan dalam judul paper sebelumnya. Dalam konteks ini, metode Decision Tree C 4.5 digunakan untuk membuat keputusan klasifikasi berdasarkan aturan-aturan yang dibangun dari data pelatihan.

### 2. Pengertian data mining

Data mining adalah proses penemuan pola atau penarikan informasi yang bermanfaat dan dapat diinterpretasikan dari suatu himpunan data yang besar, kompleks, dan terstruktur. Tujuan utama dari data mining adalah untuk mengungkap hubungan tersembunyi, tren, atau pola-pola yang mungkin sulit atau tidak mungkin ditemukan dengan menggunakan metode analisis data konvensional [10][12].

Proses data mining melibatkan beberapa langkah, (a) Pengumpulan Data: Mengumpulkan data dari berbagai sumber yang mungkin termasuk basis data, dokumen teks, log transaksi, dan sumber data lainnya. (b) Pembersihan Data (*Data Cleaning*): Membersihkan data dari *noise*, *outlier*, atau data yang tidak akurat untuk memastikan keandalan hasil analisis. (c) Seleksi Fitur (*Feature Selection*): Memilih subset fitur atau atribut yang paling relevan untuk analisis. (d) Transformasi Data (*Data Transformation*): Mengubah format atau representasi data untuk memudahkan proses analisis. (e) Pemodelan (*Modeling*): Menggunakan teknik-teknik analisis statistik atau metode machine learning untuk membangun model yang dapat mengidentifikasi pola atau hubungan dalam data. (f) Evaluasi Model (*Model Evaluation*): Mengevaluasi kinerja model yang dibangun dengan

menggunakan metrik evaluasi yang sesuai. (g) Interpretasi dan Penggunaan Hasil: Menginterpretasikan temuan dari proses data mining dan menggunakannya untuk mengambil keputusan atau mendukung pengambilan keputusan.

Data mining digunakan dalam berbagai bidang, termasuk bisnis, ilmu pengetahuan, kesehatan, keamanan, dan lainnya. Contoh aplikasi data mining meliputi segmentasi pelanggan, prediksi tren pasar, deteksi penipuan, diagnosis medis, dan banyak lagi. Data mining membantu menggali nilai tambah dari data besar dan kompleks, yang dapat memberikan wawasan yang berharga bagi organisasi atau individu.

### 3. Konsep Dasar Decision Tree

Decision Tree (Pohon Keputusan) adalah representasi grafis dari suatu keputusan atau serangkaian keputusan dan konsekuensinya. Dalam konteks data mining dan machine learning, Decision Tree digunakan sebagai model prediktif untuk mengambil keputusan berdasarkan aturan-aturan yang dihasilkan dari analisis data.

Decision Tree dapat dianalogikan dengan Decision Tree yang sebenarnya, di mana setiap simpul pada pohon mewakili suatu keputusan, cabang-cabang yang keluar dari simpul tersebut menggambarkan konsekuensi atau hasil dari keputusan tersebut, dan daun-daun pada pohon menyatakan label kelas atau nilai prediksi akhir. Decision Tree cocok digunakan untuk tugas klasifikasi dan regresi. Algoritma Decision Tree C 4.5 adalah salah satu contoh populer dari algoritma pembentukan Decision Tree yang digunakan dalam konteks klasifikasi [7].

### 4. Decision Tree C 4.5

Decision Tree C 4.5 (atau juga dikenal sebagai C 4.5) adalah algoritma pembentukan Decision Tree yang dikembangkan oleh Ross Quinlan. Algoritma ini merupakan evolusi dari pendahulunya, ID3 (Iterative Dichotomiser 3), dan dirancang untuk menangani masalah klasifikasi pada data yang kompleks dan tidak terstruktur [2].

Berikut adalah beberapa karakteristik dan konsep utama dari Decision Tree C 4.5:

1. Pemilihan Atribut: Algoritma ini menggunakan metode pengukuran kriteria seperti Information Gain untuk menentukan atribut mana yang paling baik untuk membagi data pada setiap simpul (node) dalam pohon keputusan.
2. Pemisahan Data: Setiap simpul pada Decision Tree membagi data menjadi subset berdasarkan nilai dari atribut yang dipilih. Proses ini dilakukan secara rekursif untuk membentuk struktur pohon yang lebih kompleks.
3. Pengukuran Kepentingan Atribut: Decision Tree C 4.5 mengukur seberapa baik suatu atribut dapat memisahkan kelas-kelas pada data. Informasi Gain atau Gini Index adalah metrik yang digunakan untuk menentukan atribut yang paling informatif.
4. Pruning (Pemangkasan Pohon): Untuk mencegah overfitting, Decision Tree C 4.5 dapat mengalami proses pruning, di mana cabang-cabang atau simpul-simpul yang kurang relevan atau terlalu spesifik dihapus.
5. Penanganan Data yang Hilang: C4.5 dapat mengatasi masalah missing values dengan menggunakan strategi penggantian nilai yang tepat.
6. Pengklasifikasian pada Daun (*Leaf*): Ketika mencapai daun pada pohon, keputusan klasifikasi dibuat berdasarkan mayoritas kelas pada subset data tersebut.

Decision Tree C4.5 banyak digunakan dalam berbagai bidang, termasuk data mining, machine learning, dan pengambilan keputusan. Kelebihannya meliputi kemampuan interpretabilitas yang tinggi, kemampuan menangani data kategori, serta kesesuaian dengan berbagai jenis data.

### 5. Kredit Pinjaman

Kredit adalah bentuk fasilitas keuangan di mana pihak pemberi kredit memberikan dana atau sumber daya kepada pihak peminjam dengan harapan bahwa peminjam akan mengembalikan jumlah yang dipinjam beserta bunga atau biaya lainnya dalam jangka waktu yang telah disepakati. Kredit merupakan alat penting dalam aktivitas keuangan, baik bagi individu, perusahaan, maupun lembaga keuangan [3][5][11].

Kredit dapat menjadi instrumen yang memungkinkan individu dan perusahaan untuk mencapai tujuan finansialnya, namun juga harus dikelola dengan bijak untuk menghindari risiko keuangan yang mungkin timbul [4][8].

## 6. RapidMiner

RapidMiner adalah platform open-source untuk analisis data dan penggalian pola data (data mining). Ini menyediakan berbagai alat dan fungsi untuk menjelajahi, memproses, dan menganalisis data dari berbagai sumber. RapidMiner dapat digunakan untuk tugas seperti pemrosesan data, pemodelan prediktif, pemodelan klustering, pemodelan asosiasi, dan berbagai jenis analisis data [6]. Beberapa fitur utama RapidMiner termasuk:

1. Antarmuka Pengguna Grafis: RapidMiner menawarkan antarmuka pengguna grafis yang intuitif dan mudah digunakan. Ini memungkinkan pengguna untuk membangun alur kerja analisis data tanpa harus menulis kode.
2. Libraries Algoritma: RapidMiner menyediakan berbagai algoritma machine learning dan statistik yang dapat digunakan untuk mengembangkan model prediktif dan menggali pola data.
3. Integrasi Data: RapidMiner dapat mengintegrasikan data dari berbagai sumber, termasuk basis data, file teks, dan sumber data lainnya.
4. Validasi Model: Alat validasi model memungkinkan pengguna untuk menguji dan memvalidasi keakuratan model yang dibuat.
5. Ekstensi dan Kustomisasi: RapidMiner dapat dikustomisasi dan diperluas melalui ekstensi dan plug-in.
6. Komunitas dan Dukungan: RapidMiner memiliki komunitas yang aktif yang dapat memberikan dukungan dan berbagi pengetahuan.

Platform ini digunakan di berbagai industri dan disukai oleh data scientist, analis data, dan profesional di bidang analisis data. Meskipun RapidMiner tersedia dalam versi open-source, ada juga versi komersial yang menyediakan fitur tambahan dan dukungan yang lebih mendalam.

## 3. METODE PENELITIAN

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. [1]. Metode penelitian yang dilakukan adalah metode penelitian eksperimen, dengan tahapan sebagai berikut [14] :

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/Metode yang Diusulkan (*Proposed Model/Method*)  
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, model yang diusulkan akan diuji untuk mengetahui hasil dari peraturan yang akan digunakan untuk pengambilan keputusan.
5. Evaluasi dan Validasi Hasil (*Result Evaluation and Validation*)  
Pada tahap ini, model dievaluasi untuk mengetahui seberapa akurat model yang telah dibuat.

### 3.1. Pengumpulan Data

Data yang digunakan dalam penelitian ini adalah german credit. Data ini terbuka dan dapat di unduh di <https://github.com/selva86/datasets/blob/master/GermanCredit.csv> dan

<https://data.world/uci/statlog-german-credit-data> [6] yang berasal dari Profesor Dr. Hans Hofmann Institut Statistik dan "Ökonometrie Universit" di Hamburg FB Wirtschaftswissenschaften Von-Melle-Park 5 2000 Hamburg 13. Data tersebut terdiri dari 1000 Record dengan atribut Data tersebut berjumlah 1000 record yang terdiri dari 20 atribut yaitu: C1: Status rekening giro (4 Kategori), C2: jangka waktu peminjaman, C3: sejarah kredit (5 Kategori), C4: tujuan kredit (11 Kategori), C5: jumlah kredit, C6: jumlah simpanan (5 Kategori), C7: lama bekerja (5 Kategori), C8: biaya angsuran terhadap sisa pendapatan, C9: status perkawinan dan jenis kelamin (5 Kategori), C10: status orang yang berhutang lainnya (penjamin) (3 Kategori), C11: lama bertempat tinggal, C12: properti (4 Kategori), C13: usia, C14: rencana angsuran lain (3 Kategori), C15: tempat tinggal (5 Kategori), C16: jumlah kredit sebelumnya, C17: pekerjaan (4 Kategori), C18: jumlah tanggungan, C19: telepon (2 Kategori), dan C20: tenaga kerja asing (2 Kategori). Sedangkan kategori numeriknya adalah C2: Jangka waktu peminjaman (bulan), C5: Jumlah kredit (euro), C8: Biaya angsuran terhadap sisa pendapatan (frekuensi angsuran tiap bulan), C11: Lama bertempat tinggal (tahun), C13: Usia (tahun), C16: Jumlah kredit sebelumnya (frekuensi kredit), C18: Jumlah tanggungan (jumlah orang yang ditanggung).

### 3.2. Pengolahan data awal

Data yang didapatkan dari sumber 1 merupakan data asli yang belum dilakukan perubahan supaya lebih mudah pembacaannya dan setelah melalui proses perapihan data, sumber 2 memberikan kemudahan untuk merapihkan data yang akan diproses.

### 3.3. Metode yang diusulkan

Metode yang diusulkan untuk menentukan klasifikasi adalah metode Decision Tree Algoritma C4.5, yang merupakan salah satu teknik klasifikasi yang paling populer dalam data mining. Metode Decision Tree mengubah banyak fakta menjadi Decision Tree yang berfungsi sebagai representasi aturan. Untuk mengeksplorasi data, Decision Tree berguna untuk menemukan hubungan tersembunyi antara sebuah variabel target dan berbagai variabel input potensial. Data biasanya disimpan dalam Decision Tree dalam bentuk tabel yang memiliki record dan atribut. Dalam pembentukan pohon keputusan, atribut menunjukkan parameter yang dikenal sebagai kriteria.

Dengan menggunakan software RapidMiner 5.3 untuk pembuatan analisa dan pengujian model. Tahapan Algoritma Decision Tree C4.5:

- 1) Menyiapkan *data training*
- 2) Menentukan akar dari pohon.
- 3) Hitung nilai *Gain*:

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

- 4) Ulangi langkah ke-2 hingga semua tupel terpartisi

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

- 5) Proses partisi Decision Tree akan berhenti saat semua tupel dalam node N mendapat kelas yang sama dan atau tidak ada atribut di dalam tupel yang dipartisi lagi dan atau tidak ada tupel di dalam cabang yang kosong.

### 3.4. Experiment dan pengujian metode

Metode yang telah dikembangkan dalam penelitian ini akan diterapkan pada data yang sudah diolah dari sumber 1 dan sumber 2. Penerapan Metode Decision Tree C4.5 untuk Klasifikasi penilaian kemampuan kredit melalui suatu simulasi menggunakan software RapidMiner 5.3. Data Sampel terdiri dari 20 atribut dengan jumlah data sebanyak 1000 sampel. Data Training yang digunakan adalah 70%, 75%, 80%, 85% dan 90% sedangkan data testingnya adalah 30%, 25%, 20%, 15% dan 10%. Perbedaan pada data training dan data testing untuk mengetahui akurasi tertinggi. Proses Pengujiannya adalah sebagai berikut:

### 1. Import data dari sumber data.

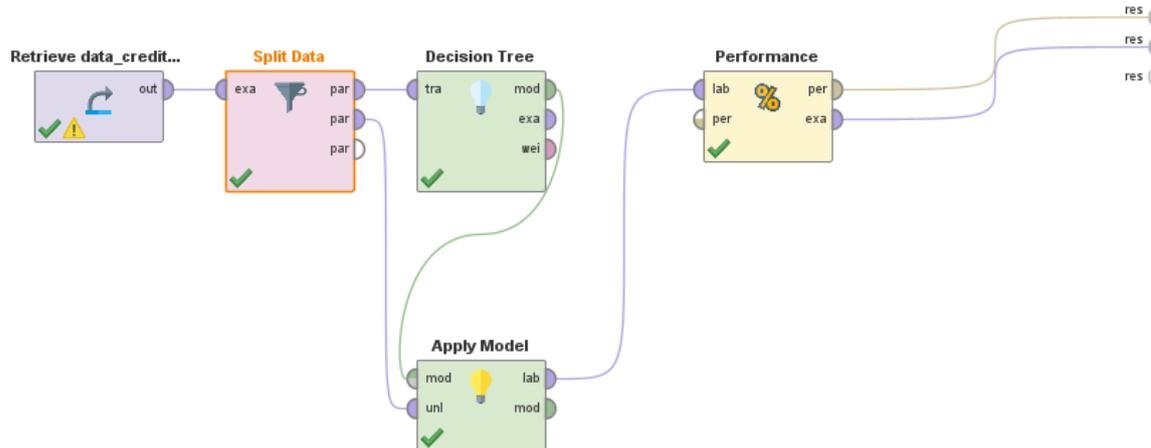
Melakukan import dari data yang sudah dilakukan pengolahan supaya dapat diproses menggunakan aplikasi rapid manner. Kemudian salah satu attribute diberikan label sebagai prediksi. berikut ini contoh dari bentuk data yang akan diproses 13 data dari 1000 data yang diproses.

...	↑	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19	C20	C21
1		A11	6	A34	A43	1169	A65	A75	4	A93	A101	4	A121	67	A143	A152	2	A173	1	A192	A201	baik
2		A12	48	A32	A43	5951	A61	A73	2	A92	A101	2	A121	22	A143	A152	1	A173	1	A191	A201	buruk
3		A14	12	A34	A46	2096	A61	A74	2	A93	A101	3	A121	49	A143	A152	1	A172	2	A191	A201	baik
4		A11	42	A32	A42	7882	A61	A74	2	A93	A103	4	A122	45	A143	A153	1	A173	2	A191	A201	baik
5		A11	24	A33	A40	4870	A61	A73	3	A93	A101	4	A124	53	A143	A153	2	A173	2	A191	A201	buruk
6		A14	36	A32	A46	9055	A65	A73	2	A93	A101	4	A124	35	A143	A153	1	A172	2	A192	A201	baik
7		A14	24	A32	A42	2835	A63	A75	3	A93	A101	4	A122	53	A143	A152	1	A173	1	A191	A201	baik
8		A12	36	A32	A41	6948	A61	A73	2	A93	A101	2	A123	35	A143	A151	1	A174	1	A192	A201	baik
9		A14	12	A32	A43	3059	A64	A74	2	A91	A101	4	A121	61	A143	A152	1	A172	1	A191	A201	baik
10		A12	30	A34	A40	5234	A61	A71	4	A94	A101	2	A123	28	A143	A152	2	A174	1	A191	A201	buruk
11		A12	12	A32	A40	1295	A61	A72	3	A92	A101	1	A123	25	A143	A151	1	A173	1	A191	A201	buruk
12		A11	48	A32	A49	4308	A61	A72	3	A92	A101	4	A122	24	A143	A151	1	A173	1	A191	A201	buruk
13		A12	12	A32	A43	1567	A61	A73	1	A92	A101	1	A123	22	A143	A152	1	A173	1	A192	A201	baik

**Gambar 1.** Sumber Data

### 2. Membuat Desain Data Proses

Pada tahap ini penulis membuat desain untuk proses *experiment* dan pengujian dengan membuat desain seperti gambar 2.



**Gambar 2.** Desain Proses (*Rapid Maner*)

Retrieve data\_Credit adalah data yang telah diimport dari sumber data yang akan diproses dalam penelitian ini yang terdiri dari 1000 data yang sudah dilabeli salah satu atributnya sebagai prediksi. Split data adalah pembagian data berapa persen untuk data training dan berapa persen untuk data testing. Decission Tree dan Apply Model adalah proses dimana algoritma Decision Tree C 4.5 dikompilasi. Performance adalah untuk mengetahui hasil evaluasi Performansinya.

### 3. Running Program dan Evaluasi Performance

Pada tahap ini program dijalankan untuk menghasilkan output performance yang dihasilkan dengan melakukan perubahan prosentase split data.

### 3.5 Evaluasi dan Validasi Hasil

Evaluasi dilakukan dengan menganalisa hasil klasifikasi dari penggunaan metode algoritma Decision Tree C4.5, yaitu membandingkan nilai Akurasi, nilai *Precision*, dan nilai *Recall*. Perhitungan akurasi dengan tabel confusion matrix adalah sebagai berikut:

$$\text{Akurasi} = \frac{F_{11} + F_{00}}{F_{11} + F_{10} + F_{01} + F_{00}} \quad (1)$$

Penjelasan tentang pengukuran *Precision* dan *Recall* dapat dilihat pada Tabel 1.

**Tabel 1.** Perhitungan *Precision* dan *Recall*

	Relevant	Not Relevant
Retrieved	A	B
Not Retrieved	C	D

Pada evaluasi ini dengan melakukan perubahan prosentase terhadap *data training* dan *data testing* pada sampel 1000 data input. Pengukuran dilakukan dengan melakukan perubahan data training dan data testing, kemudian melihat performansi dari perubahan data tersebut untuk mengetahui akurasi tertinggi dari perubahan data yang dilakukan. Disajikan pada Tabel 2.

**Tabel 2.** Perubahan *Data Testing* dan *Data Training*

No	Data Training	Data Testing	Total
1	70%	30%	100%
2	75%	25%	100%
3	80%	20%	100%
4	90%	10%	100%

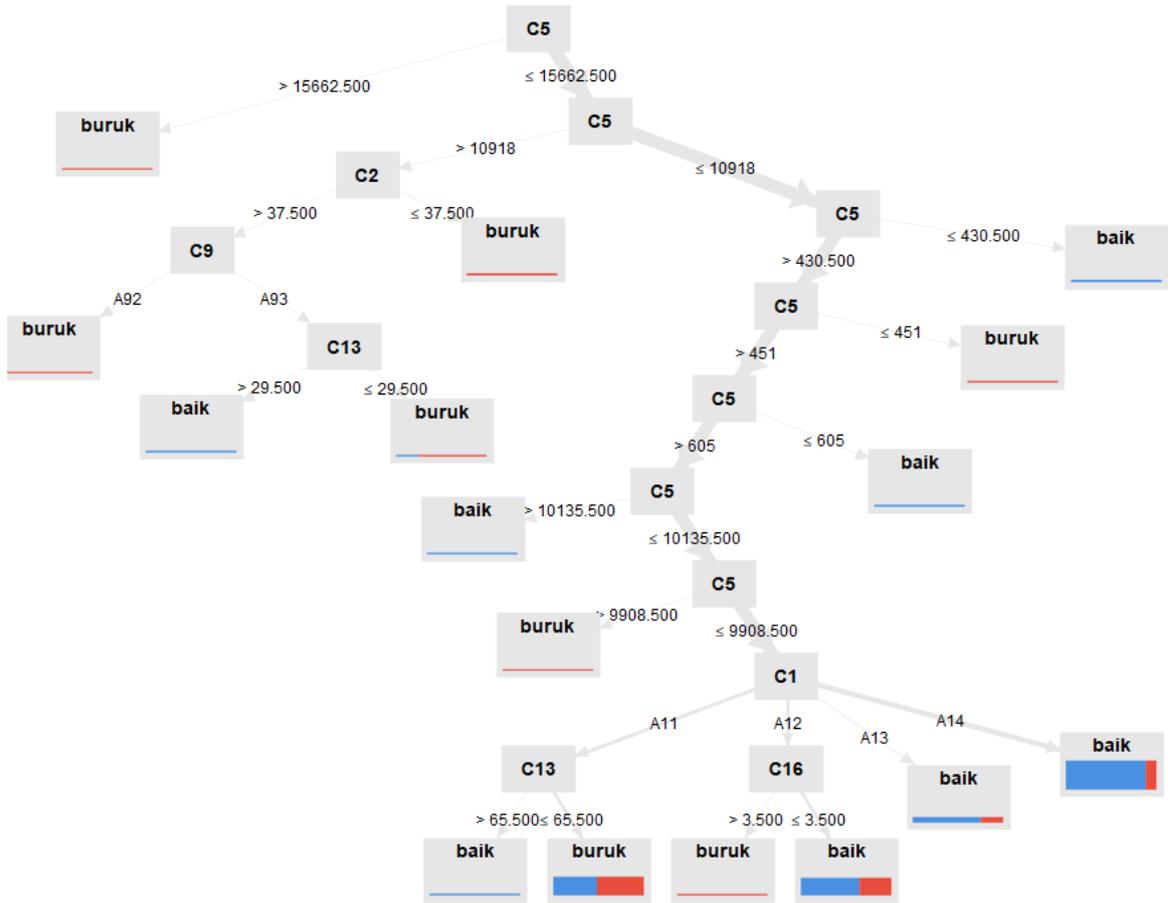
## 4. HASIL DAN PEMBAHASAN

Pada tahap akhir penerapan algoritma Decision Tree C 4.5 berdasarkan pengujian yang dilakukan dengan menggunakan RapidMiner memberikan hasil Decision Tree seperti terlihat pada Gambar 3. Notasi C menunjukkan atribut yang sudah didefinisikan di awal, sedang baik dan buruk adalah target prediksi, yaitu atribut yang sudah diberi label pada pendefinisian data sebelum diinputkan pada proses desain yang telah dibuat.

**Tabel 3.** Performansi Sistem

No	Data Training (%)	Data Testing (%)	Recall (TP)	Recall (TN)	Precision (PP)	Precision (PN)	Akurasi
1	70	30	76,19	56,67	80,40	50,05	70,33
2	75	25	98,86	8,00	71,49	75,00	71,60
3	80	20	77,14	55,0	80,00	50,77	70,50
4	85	15	79,05	51,11	79,05	51,11	70,67
5	90	10	70,00	56,67	79,03	44,74	66,00

Dari tabel 2 dapat dilihat akurasi paling baik terjadi pada pembagian data training 70% dan data testing 25% yaitu akurasinya 71,60.



Gambar 3. Pohon Keputusan

Gambar 3 diatas merupakan Decision Tree yang dihasilkan dari *rule* yang dapat dilihat pada *text view* pada Gambar 4.

**Tree**

```

C5 > 15662.500: buruk {baik=0, buruk=2}
C5 ≤ 15662.500
| C5 > 14248.500: buruk {baik=1, buruk=3}
| C5 ≤ 14248.500
| | C5 > 430.500
| | | C5 > 451
| | | | C5 > 634
| | | | | C5 > 663
| | | | | | C1 = A11
| | | | | | | C13 > 65.500: baik {baik=4, buruk=0}
| | | | | | | C13 ≤ 65.500: buruk {baik=67, buruk=76}
| | | | | | | C1 = A12
| | | | | | | C5 > 12391.500: buruk {baik=0, buruk=3}
| | | | | | | C5 ≤ 12391.500
| | | | | | | | C16 > 3.500: buruk {baik=0, buruk=2}
| | | | | | | | C16 ≤ 3.500: baik {baik=95, buruk=52}
| | | | | | | C1 = A13
| | | | | | | | C5 > 1036.500: baik {baik=27, buruk=6}
| | | | | | | | C5 ≤ 1036.500: buruk {baik=1, buruk=3}
| | | | | | | C1 = A14
| | | | | | | | C5 > 11232.500: buruk {baik=1, buruk=2}
| | | | | | | | C5 ≤ 11232.500: baik {baik=206, buruk=26}
| | | | | | | C5 ≤ 663: buruk {baik=0, buruk=2}
| | | | | | | C5 ≤ 634: baik {baik=9, buruk=0}
| | | | | | | C5 ≤ 451: buruk {baik=0, buruk=3}
| | | | | | | C5 ≤ 430.500: baik {baik=9, buruk=0}
    
```

Gambar 4. Text View Decision Tree C 4.5

Berdasarkan perubahan yang dilakukan terhadap *data training* dan *data testing* diperoleh *recall*, *precision* dan akurasi sesuai yang tertera pada Tabel 3.

## KESIMPULAN

Penelitian eksperimental ini mengikuti tahapan pengumpulan data, praproses, usulan model, pengujian, dan evaluasi hasil. Data kredit Jerman dengan 1000 catatan dan 20 atribut digunakan. Algoritma Decision Tree C4.5 diterapkan menggunakan RapidMiner, dengan data dibagi menjadi set pelatihan dan pengujian untuk menentukan akurasi tertinggi. Algoritma Decision Tree C 4.5 yang diterapkan di RapidMiner menghasilkan Decision Tree dan kumpulan aturan. Akurasi terbaik dicapai dengan 75% data latih dan 25% data uji, sehingga menghasilkan akurasi 71,60%. Metode Decision Tree C 4.5 secara efektif mengklasifikasikan data kredit, memberikan wawasan berharga untuk evaluasi risiko kredit. Penelitian ini berkontribusi pada pengembangan model penerimaan kredit yang andal, bermanfaat bagi lembaga keuangan dan masyarakat.

## DAFTAR PUSTAKA

- [1] Berndtsson M, Hansson J, Olsson B, Lundell B. (2008). Thesis Projects: A Guide for Students in Computer Science and Information Systems. 165 p
- [2] Dewi, I. G. A. M. P., Parwita, W. G. S., & Setiawan, I. M. D. (2021). Algoritma Decision Tree untuk Klasifikasi Calon Debitur LPD Desa Adat Anggungan. *Jurnal Krisnadana*, 1(1), 23-36.
- [3] Handayani, N., Wahyono, H., Trianto, J., & Permana, D. S. (2021). Prediksi Tingkat Risiko Kredit dengan Data Mining Menggunakan Algoritma Decision Tree C. 45. *JURIKOM (Jurnal Riset Komputer)*, 8(6), 198-204.
- [4] Kholifah, A. N., & Insani, N. (2016). Analisis Klasifikasi Pada Nasabah Kredit Koperasi X Menggunakan Decision Tree C4. 5 Dan Naive Bayes. *Jurnal Kajian dan Terapan Matematika*, 5(6).
- [5] Pratama, A. Z., Kurniawati, L., Larbona, S., & Haryanti, T. (2019). Algoritma C4. 5 Untuk Klasifikasi Nasabah Dalam Memprediksi Kredit Macet. *Information System For Educators And Professionals: Journal of Information System*, 3(2), 121-130.
- [6] PUTRA, R. H. (2016). RANCANG BANGUN SISTEM PENDUKUNG KEPUTUSAN KLASIFIKASI KELAS RISIKO KREDIT MENGGUNAKAN METODE NAÏVE BAYES (Doctoral dissertation, UNIVERSITAS AIRLANGGA).
- [7] RIDWAN, F. A. (2016). SISTEM PENDUKUNG KEPUTUSAN PEMBERIAN KELAYAKAN PENGAJUAN KREDIT SEPEDA MOTOR MENGGUNAKAN METODE DECISION TREE C 4.5 (Doctoral dissertation, Universitas Muhammadiyah Gresik).
- [8] Rosihan, R., Fhadli, M., & Usman, A. A. H. (2023). Klasifikasi Kelayakan Pemberian Kredit Menggunakan Metode Decision Tree dengan Seleksi Fitur (Studi Kasus: PT. Adira Finance Cabang Kota Ternate). *Jurnal Pendidikan Tambusai*, 7(3), 21517-21524.
- [9] Rusito, R., & Firmansyah, M. (2016). Implementasi Metode Decision Tree Dan Algoritma C4. 5 Untuk Klasifikasi Data Nasabah Bank. *Jurnal Ilmiah Infokam*, 12(2).
- [10] Sari, R. D. I., & Sindunata, Y. (2014). PENERAPAN DATA MINING UNTUK ANALISA POLA PERILAKU NASABAH DALAM PENGKREDITAN MENGGUNAKAN METODE C. 45 STUDI KASUS PADA KSU INSAN KAMIL DEMAK. *Jurnal Ilmiah Teknologi Informasi Asia*, 8(2), 10-16.
- [11] Setiadi, B., & Lareno, B. (2015). Penerapan Algoritma Decision Tree Untuk Penilaian Agunan Pengajuan Kredit. *Proceedings Konferensi Nasional Sistem dan Informatika (KNS&I)*.
- [12] Suherman, S., Maulana, D., & Mustikaningtyas, V. (2022). Penerapan Data Mining Untuk Klasifikasi Kepuasan Pelanggan Transportasi Online (Ojek Online) Menggunakan Algoritma C. 4.5. *Prosiding Sains dan Teknologi*, 1(1), 165-175.
- [13] Sunarti, S. (2021). Klasifikasi Penentuan Kelayakan Pemberian Pinjaman Pada Koperasi Karyawan Menggunakan Algoritma C4. 5. *JOINS (Journal of Information System)*, 6(1), 1-8.
- [14] Hijriana, N. and R. Muttaqin (2016). "Penerapan Metode Decision Tree Algoritma C4. 5 Untuk Klasifikasi Mahasiswa Berprestasi." *AL ULUM: JURNAL SAINS DAN TEKNOLOGI* 2(1): 39-43.