

PENERAPAN METODE *DECISION TREE* DALAM MEMPREDIKSI PERSETUJUAN KREDIT BANK

**Disusun guna memenuhi tugas mata kuliah
Data Warehouse & Business Intelligence**



Disusun Oleh:

Kelompok 7

- | | |
|--------------------------|----------|
| 1. Affreza Dinova | 19248132 |
| 2. Nalendro Adil Satyadi | 19248110 |
| 3. Ivone Riza Pangesti | 19248078 |
| 4. Fiki Nur Alifah | 19248128 |
| 5. Sabrina Viola Sari | 19248088 |

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK & INFORMATIKA
UNIVERSITAS BINA SARANA INFORMATIKA
2024**

KATA PENGANTAR

Puji syukur kepada Tuhan Yang Maha Esa karena rahmat, kasih setia dan penyertaannya yang selalu kita rasakan setiap hari. Puji syukur kepada atas kasih karunia dan keselamatan, serta atas tuntunan, penyertaan, rahmat dan akal budi sehingga kami dapat menyelesaikan Tugas Proyek Pada Mata Kuliah Data Warehouse dan Business Intelligence ” Penerapan Metode *Decision Tree* Dalam Memprediksi Persetujuan Kredit Bank ”.

Pada kesempatan kali ini kami ingin mengucapkan terimakasih kepada Bapak Hafdiarsya Saiyar, ST, M.Kom selaku dosen pada Tugas Proyek Pada Mata Kuliah Data Warehouse dan Business Intelligence. Kami sangat berharap semoga makalah ini dapat membantu dan menambah pengetahuan bagi pembaca, kami pun menyadari masih banyak kekurangan dalam penyusunan maupun kata-kata yang kami gunakan dalam makalah ini. Untuk itu kami sangat mengharapkan kritik dan saran yang membangun dari pembaca demi kesempurnaan makalah ini.

Banyumas, 29 November 2024

Penyusun

DAFTAR ISI

| | |
|---|-----|
| HALAMAN JUDUL | i |
| KATA PENGANTAR..... | ii |
| DAFTAR ISI | iii |
| BAB I PENDAHULUAN | 1 |
| 1.1 Latar Belakang Masalah | 1 |
| 1.2 Tujuan Dan Manfaat | 1 |
| 1.3 Metode Penelitian | 2 |
| 1.3.1 Pengumpulan Data..... | 2 |
| 1.4 Ruang Lingkup | 3 |
| BAB II LANDASAN TEORI..... | 4 |
| 2.1 Pengertian Data Warehouse | 4 |
| 2.2 Pengertian Business Intelligence | 4 |
| 2.3 Pengertian Rapid Miner..... | 4 |
| 2.4 Pengertian Data Warehouse Pada Bidang Perbankan | 5 |
| 2.5 Pengertian Sistem Informasi Akademik..... | 5 |
| BAB III PEMBAHASAN | 6 |
| 3.1 Metode Pengumpulan Data | 6 |
| 3.2 Analisa Data | 6 |
| 3.2.1 Analisis Data | 6 |
| 3.2.2 Proses Analisis Data | 7 |
| 3.3 Pengolahan Data | 7 |
| 3.4 Hasil Penelitian..... | 10 |
| BAB IV PENUTUP..... | 13 |
| 4.1 Kesimpulan..... | 13 |
| 4.2 Saran | 13 |
| DAFTAR PUSTAKA..... | 14 |

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Teknologi informasi yang terus berkembang selalu dituntut untuk dapat memenuhi berbagai kebutuhan di segala bidang kehidupan. Teknologi informasi juga dituntut untuk memudahkan manusia dalam beraktivitas, seperti pengambilan keputusan, memecahkan masalah komputasi yang kompleks, mengakses informasi dan mengirim data, serta beberapa kebutuhan manusia yang lainnya. Salah satunya dalam bidang perbankan, yaitu implementasi *data warehouse* dalam bank harus memperhatikan pemilihan platform dan keamanan data, karena keamanan data juga sangat penting untuk diperhatikan mengingat bank menyimpan data personal nasabah yang meliputi data diri dan transaksi.

Dalam dunia perbankan, penerapan business intelligence merupakan kunci sukses dalam mengefisiensikan dan mengefektifkan kegiatan bisnis utama dengan kemampuan dalam mendapatkan, mengelola dan menganalisa data nasabah serta layanan dalam jumlah yang besar. pemanfaatan teknologi ini memungkinkan bank untuk memberikan layanan yang lebih baik kepada nasabah, seperti personal layanan, prediksi kebutuhan nasabah, hingga mendeteksi potensi resiko yang dapat mempengaruhi operasional.

Namun, penerapan teknologi ini tidak lepas dari tantangan yang harus dihadapi. Bank harus dapat memastikan bahwa sistem data warehouse dan business intelligence yang digunakan memiliki tingkat keamanan tinggi untuk mencegah ancaman kebocoran data. Oleh karena itu, pemanfaatan data warehouse dan business intelligence yang didukung teknologi handal menjadi kebutuhan yang dapat diandalkan.

1.2 Tujuan Dan Manfaat

Berdasarkan latar belakang yang telah dijelaskan, berikut adalah tujuan dan manfaat :

Tujuan :

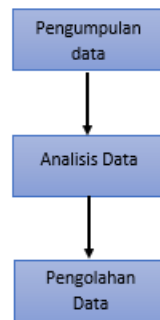
- a. Untuk meningkatkan efisiensi pengelolaan data

- b. Untuk meningkatkan keamanan data nasabah
- c. Agar dapat mendukung pengambilan keputusan yang strategis
- d. Untuk meningkatkan kualitas layanan nasabah

Manfaat :

- a. Efisiensi operasional
- b. Keamanan informasi
- c. Pengambilan keputusan yang tepat

1.3 Metode Penelitian



Sumber: (Rofiani et al., 2024)

Gambar I.1 Metode Penelitian

1.3.1 Pengumpulan Data

- a. Sumber Data : Dataset ini diambil dari situs web UCI Repository dan terdiri dari dataset yang relevan dengan kredit bank.
- b. Atribut yang Tersedia: Dataset ini memiliki 14 atribut yang berbeda, yaitu:
 1. Nama : Nama Subjek dalam dataset.
 2. Jenis Kelamin : Jenis kelamin subjek, misalnya, pria atau wanita.
 3. Umur : Umur Subjek dalam dataset.

4. Jumlah Pinjaman : Informasi tentang Jumlah Pinjaman.
5. Jumlah Angsuran Perbulan : Informasi Tentang Jumlah Angsuran Perbulan
6. Tipe Pinjaman : Informasi Tentang Tipe Pinjaman Subjek..
7. Jenis Pinjaman : Informasi Tentang Jenis Pinjaman Subjek..
8. BI Sektor Ekonomi : Informasi Tentang Kode Sektor Ekonomi Subjek.
9. BI Golongan Debitur : Informasi Tentang Kode Golongan Debitur Subjek.
10. BI Golongan Penjamin : Informasi Tentang Kode Golongan Penjamin.
11. Saldo Nominatif : Informasi Tentang Saldo Nominatif Subjek.
12. Tunggakan Pokok : Informasi Tentang Tunggakan Pokok Subjek.
13. Tunggakan Bunga : Informasi Tentang Tunggakan Bunga Subjek.
14. Status Kredit : Status Kredit Subjek (Lancar / Macet).

1.4 Ruang Lingkup

Penelitian difokuskan pada bidang teknologi informasi dalam sektor perbankan, khususnya penerapan :

- a. Data warehouse sebagai sistem untuk menyimpan dan mengelola data dalam jumlah besar
- b. Business intelligence (BI) sebagai alat analisis untuk mendukung pengambilan keputusan strategis dan peningkatan kualitas layanan.

BAB II

LANDASAN TEORI

2.1 Pengertian Data Warehouse

Data Warehouse adalah metode untuk mengintegrasikan dan menggabungkan data dalam format multidimensi. Proses ini mencakup pembersihan, penyatuan, dan transformasi data, sehingga berfungsi sebagai tahap praproses yang penting dalam mendukung aktivitas data mining. (Wirdasari & Calam, 2011)

2.2 Pengertian Business Intelligence

Business Intelligence adalah serangkaian proses yang bertujuan untuk mengumpulkan dan menganalisis data organisasi guna mendukung pengambilan keputusan secara optimal. Selain itu, Business Intelligence juga berperan dalam memprediksi serta mengevaluasi dampak dari keputusan yang diambil oleh organisasi. Awalnya, Business Intelligence dirancang untuk membantu manajemen dalam membuat keputusan strategis, tetapi juga dapat digunakan dalam mendukung keputusan taktis di tingkat organisasi. (Pratasik, 2019)

2.3 Pengertian Rapid Miner

Menurut S. Widaningsih RapidMiner adalah platform perangkat lunak berbasis Java yang dirancang untuk analisis data, kompatibel dengan berbagai sistem operasi. Platform ini menyediakan alat analisis data yang lengkap dan mudah digunakan, bahkan tanpa keahlian teknis mendalam. Fitur utamanya meliputi pemodelan prediktif, analisis statistik, serta penambangan data, termasuk analisis teks. Dengan antarmuka yang intuitif, pengguna dapat melakukan berbagai tugas seperti pengolahan data, pembuatan model prediktif, klusterisasi, analisis teks, dan visualisasi data. RapidMiner digunakan di berbagai industri, seperti keuangan, kesehatan, e-commerce, dan pemasaran, untuk memberikan wawasan berbasis data yang mendukung pengambilan keputusan. Platform ini tersedia dalam versi open source yang menawarkan akses gratis ke banyak fitur dasar, serta didukung oleh komunitas pengguna yang aktif, yang memungkinkan pertukaran pengalaman dan solusi terkait analisis data. (Rofiani et al., 2024)

2.4 Pengertian Data Warehouse Pada Bidang Perbankan

Data warehouse dalam bidang perbankan adalah sistem terpusat yang dirancang untuk mengumpulkan, menyimpan, dan mengelola data dari berbagai sumber internal maupun eksternal, sehingga memungkinkan analisis data yang terstruktur dan mendalam. Data warehouse membantu bank mengintegrasikan informasi dari berbagai sistem operasional, seperti data transaksi, informasi pelanggan, risiko, hingga laporan keuangan, untuk mendukung pengambilan keputusan strategis.

2.5 Pengertian Sistem Informasi Akademik

Sistem Informasi Akademik (SIKAD) adalah platform yang berfungsi untuk menyimpan dan mengelola aktivitas akademik secara daring. Tujuan utamanya adalah mengintegrasikan berbagai informasi agar data selalu terkini dan dapat diakses secara real-time. SIKAD memudahkan pengelolaan data akademik secara efisien dan cepat, serta menjadi pusat informasi yang digunakan untuk memantau perkembangan studi mahasiswa. (Arde et al., 2021)

BAB III

PEMBAHASAN

3.1 Metode Pengumpulan Data

Tahap Pengumpulan data menggunakan sumber data sekunder, dengan data yang dipergunakan dalam penelitian ini yaitu diperoleh dari platform public dataset UCI Repository dengan alamat <https://archive.ics.uci.edu>. Dataset terdapat dua macam yaitu dataset private dan dataset publik. Private dataset yaitu dataset yang dapat diambil dari organisasi yang kita jadikan tempat atau objek penelitian, sedangkan dataset publik adalah dataset yang dapat diambil dari repository publik yang telah disepakati oleh para peneliti. (Melani & Sulastri, 2023)

3.2 Analisa Data

3.2.1 Analisis Data

Berikut alur analisis data yang dilakukan (Rofiani et al., 2024):

1. Tujuan Utama: Tahap awal analisis data ini bertujuan untuk memahami dataset secara menyeluruh sebelum memasuki proses pengolahan. Langkah ini memastikan data siap digunakan atau diuji dengan algoritma C4.5.
2. Identifikasi Data Duplikat: Analisis bertujuan mendeteksi dan menghapus data yang terduplikasi, karena keberadaan data ganda dapat mempengaruhi akurasi analisis dan hasil akhir.
3. Validasi Konsistensi Data: Data diperiksa untuk memastikan konsistensi dalam format dan nilai. Hal ini mencakup pengecekan apakah setiap entri sesuai dengan aturan dan tipe data yang ditentukan.
4. Perbaikan Kesalahan Data: Kesalahan atau inkonsistensi dalam data diperbaiki dengan melakukan tindakan seperti memperbaiki format tanggal, menangani data yang hilang, dan melakukan transformasi agar sesuai dengan kebutuhan algoritma C4.5.
5. Kesesuaian dengan Algoritma C4.5: Dataset dianalisis untuk memastikan data memenuhi syarat yang dibutuhkan algoritma C4.5, termasuk tipe data numerik, kategori, serta kecocokan dengan karakteristik algoritma.

3.2.2 Proses Analisis Data

Berikut alur proses analisis data yang dilakukan(Rofiani et al., 2024):

1. **Alat dan Teknik yang Digunakan:** Proses analisis data dapat mencakup penggunaan berbagai perangkat atau metode, seperti software statistik, pemrograman untuk membersihkan data, serta visualisasi data guna memperoleh pemahaman yang lebih mendalam.
2. **Validasi dan Verifikasi:** Setelah analisis selesai, data diperiksa kembali untuk memastikan bahwa semua isu yang ditemukan sebelumnya telah diperbaiki, sehingga dataset siap digunakan pada tahap berikutnya, yaitu pengolahan data dan penerapan algoritma C4.5.

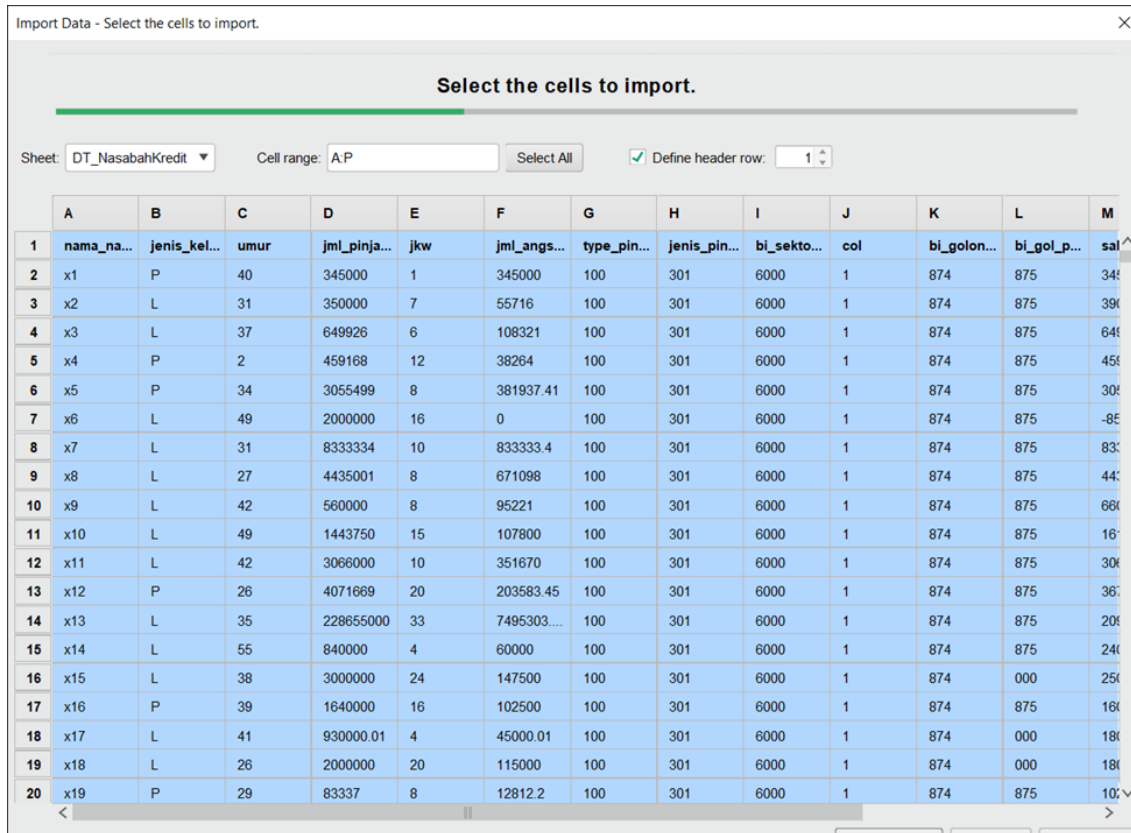
Proses analisis data ini memastikan bahwa dataset yang akan digunakan untuk pengembangan model menggunakan algoritma C4.5 telah melewati tahap pengujian dan pembersihan yang diperlukan untuk meminimalkan potensi kesalahan dan memastikan keandalan hasil analisis.(Wahid et al., 2023)

3.3 Pengolahan Data

Tahap pengolahan data dalam metode penelitian adalah proses sistematis yang mencakup serangkaian langkah penting untuk mempersiapkan data sebelum dilakukan analisis. Pengolahan dataset yang merupakan tidak lanjut dari analisis data, yang diterapkan dalam algoritma C4.5 terhadap prediksi persetujuan kredit bank yang akan menggunakan algoritma C 45 dan tools Rapidminer.(Rofiani et al., 2024)

a. Import Dataset

Data persetujuan kredit bank yang diambil dari public dataset UCI Repository. Lalu import dataset yang sudah disiapkan kedalam Rapidminer.



Gambar III. 1 *Import Dataset*

b. Cleaning Dataset

Proses pembersihan data dilakukan untuk mengatasi data yang tidak lengkap, tidak akurat, atau tidak relevan. Ini melibatkan identifikasi dan penanganan nilai yang hilang, kesalahan entri, atau ketidakkonsistenan dalam data untuk memastikan kebersihan dan keakuratan sebelum analisis.

c. Transformasi Dataset

Data sering kali memiliki format yang berbeda-beda. Tahap transformasi ini melibatkan normalisasi, pengkodean ulang variabel, atau pengubahan format data ke bentuk yang konsisten untuk memungkinkan analisis yang efektif.

d. Set Role

Set Role digunakan untuk menetapkan peran atribut dalam dataset, seperti menentukan Label pada dataset.

Import Data - Format your columns.

Format your columns.

☐ Replace errors with missing values ⓘ

| aktor_ek... r | col integer | bi_golongan... integer | bi_gol_penja... integer | saldo_nomi... real | tunggakan_... real | tunggakan_... real | status kredit polynomial label |
|------------------|----------------|---------------------------|----------------------------|-----------------------|-----------------------|-----------------------|--------------------------------------|
| 1 | 1 | 874 | 875 | 345000.000 | 345000.000 | 0.000 | MACET |
| 2 | 1 | 874 | 875 | 390000.000 | 111428.000 | 0.000 | MACET |
| 3 | 1 | 874 | 875 | 649926.000 | 216842.000 | 0.000 | MACET |
| 4 | 1 | 874 | 875 | 459168.000 | 382640.000 | 0.000 | MACET |
| 5 | 1 | 874 | 875 | 3055499.000 | 1527749.480 | 0.000 | MACET |
| 6 | 1 | 874 | 875 | -85000.000 | 0.000 | 0.000 | LANCAR |
| 7 | 1 | 874 | 875 | 8333334.000 | 3333333.600 | 0.000 | MACET |
| 8 | 1 | 874 | 875 | 4435001.000 | 0.000 | 0.000 | LANCAR |
| 9 | 1 | 874 | 875 | 660800.000 | 100800.000 | 0.000 | MACET |
| 10 | 1 | 874 | 875 | 1617000.000 | 1078000.000 | 0.000 | MACET |
| 11 | 1 | 874 | 875 | 3066000.000 | 613200.000 | 90140.000 | MACET |
| 12 | 1 | 874 | 875 | 3671669.000 | 1228667.600 | 0.000 | MACET |
| 13 | 1 | 874 | 875 | 209404092.000 | 91612122.240 | 7362732.000 | MACET |
| 14 | 1 | 874 | 875 | 240000.000 | 240000.000 | 0.000 | MACET |
| 15 | 1 | 874 | 0 | 2500000.000 | 875000.000 | 157500.000 | MACET |
| 16 | 1 | 874 | 875 | 1600000.000 | 985000.000 | 0.000 | MACET |
| 17 | 1 | 874 | 0 | 180000.010 | 180000.010 | 0.000 | MACET |
| 18 | 1 | 874 | 0 | 1800000.000 | 700000.000 | 105000.000 | MACET |

no problems.

Gambar III. 2 Set Role

e. Split Data

Split data digunakan untuk membagi data menjadi dua bagian yaitu dataset training 70% dan dataset testing 30%.

f. Modeling

Menggunakan Decision Tree untuk memodelkan data.

g. Apply Model

Pengujian model dilakukan menggunakan operator Apply Model untuk menerapkan model yang telah dilatih pada data yang belum dikenali dan menguji model berdasarkan koneksi operator Split Data, serta dilengkapi dengan operator Performance (Classification) untuk mengevaluasi kinerja pola yang dihasilkan.

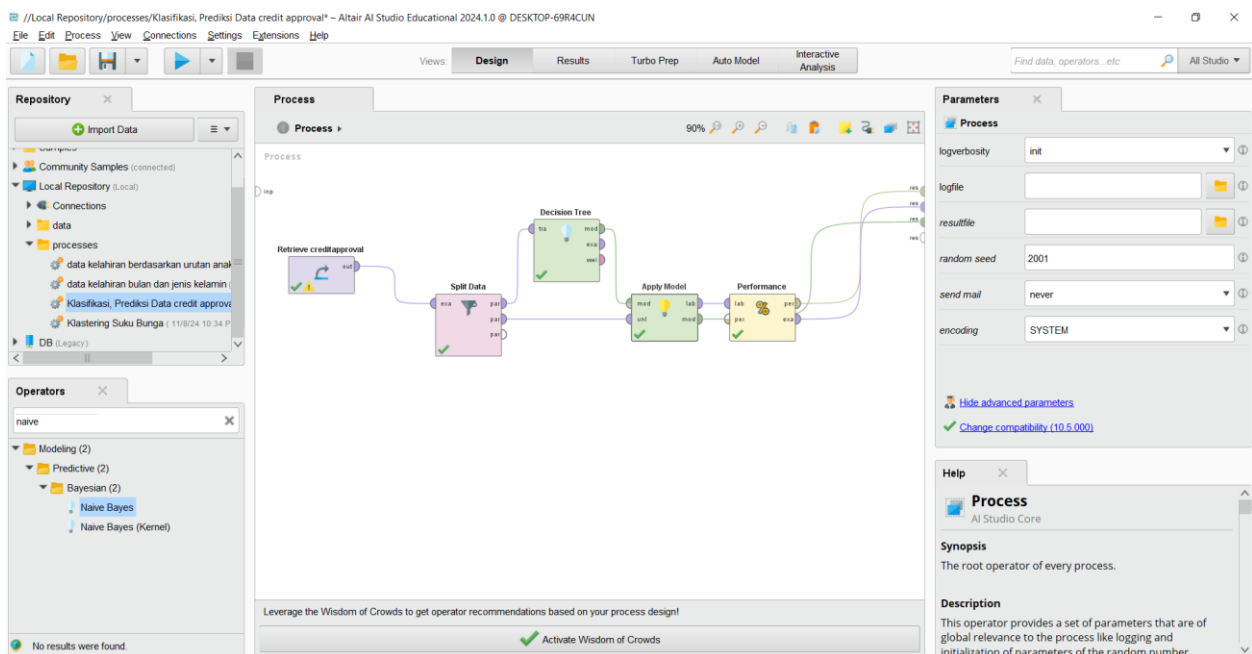
h. Evaluation

Evaluasi model dengan Decision Tree bertujuan untuk mengetahui apakah model mampu membuat prediksi dengan akurasi tinggi dan menangani ketidakseimbangan kelas menggunakan berbagai metrik seperti akurasi, precision, dan recall.

3.4 Hasil Penelitian

Proses modeling bertujuan untuk membentuk pola serta menguji model *Decision Tree* dengan membagi *dataset* menjadi data *training* dan data *testing* menggunakan rasio 70:30, di mana 70% digunakan sebagai data *training* dan 30% sebagai data *testing*. Pembagian ini dilakukan dengan menggunakan operator *Split Data* dalam RapidMiner dan mengatur parameter *split ratio* menjadi 0.7.

Setelah pembagian data, langkah berikutnya adalah membentuk model menggunakan algoritma *Decision Tree*. Tambahkan operator *Decision Tree* ke dalam proses, lalu hubungkan dengan data training dari *Split Data* untuk melatih model. Model yang telah terbentuk kemudian diuji menggunakan operator *Apply Model*, yang berfungsi untuk menerapkan model pada data testing dari *Split Data*.



Langkah terakhir adalah mengevaluasi kinerja model dengan menambahkan operator *Performance (Classification)*. Operator ini akan menghitung metrik seperti **akurasi**, **precision**, dan **recall** untuk menilai performa model dalam memprediksi data testing.

accuracy: 94.98%

| | true MACET | true LANCAR | class precision |
|--------------|------------|-------------|-----------------|
| pred. MACET | 175 | 4 | 97.77% |
| pred. LANCAR | 9 | 71 | 88.75% |
| class recall | 95.11% | 94.67% | |

Hasil perhitungan **akurasi**, baik secara manual maupun menggunakan RapidMiner, menunjukkan angka **94.98%**, yang menunjukkan bahwa model **Decision Tree** mampu memprediksi potensi penjualan dengan tingkat kesalahan yang rendah. Angka ini mengindikasikan bahwa model memiliki kinerja yang baik dalam memprediksi data, dengan hanya sekitar 5.02% kesalahan pada prediksi yang dilakukan.

precision: 88.75% (positive class: LANCAR)

| | true MACET | true LANCAR | class precision |
|--------------|------------|-------------|-----------------|
| pred. MACET | 175 | 4 | 97.77% |
| pred. LANCAR | 9 | 71 | 88.75% |
| class recall | 95.11% | 94.67% | |

Berdasarkan hasil evaluasi model, perhitungan performa *precision* sebesar **94.67%** untuk **potensi Kredit Macet** menunjukkan bahwa model memiliki kinerja yang baik, dengan hanya sedikit kesalahan dalam memprediksi kelas ini. Selain itu, *precision* **97.77%** menunjukkan bahwa model memiliki kinerja yang sangat baik dengan hanya sedikit kesalahan (terdapat **4 data** yang salah diprediksi).

Sementara itu, untuk potensi **Kredit Lancar**, hasil menunjukkan *precision* **88.75%**, yang menunjukkan bahwa model memiliki kinerja yang cukup baik meskipun ada beberapa kesalahan (terdapat **9 data** yang salah diprediksi). Meskipun model mampu memprediksi

dengan baik untuk sebagian besar data, kesalahan dalam kelas potensi **Kredit Lancar** perlu dianalisis lebih lanjut untuk meningkatkan akurasi prediksi pada kelas tersebut.

recall: 94.67% (positive class: LANCAR)

| | true MACET | true LANCAR | class precision |
|--------------|------------|-------------|-----------------|
| pred. MACET | 175 | 4 | 97.77% |
| pred. LANCAR | 9 | 71 | 88.75% |
| class recall | 95.11% | 94.67% | |

Berdasarkan hasil evaluasi model, perhitungan **recall** menunjukkan bahwa hampir seluruh data diprediksi dengan baik, dengan kesalahan prediksi sebanyak **9 data** untuk kelas **kredit lancar** dan **4 data** untuk kelas **kredit macet**. Meskipun **recall** untuk prediksi **kredit lancar** sedikit lebih rendah dibandingkan dengan **kredit macet**, model tetap efektif dalam mengenali dan memprediksi kedua kelas dengan cukup baik. Hal ini menunjukkan bahwa model mampu menangkap sebagian besar data yang relevan dengan tingkat kesalahan yang minimal.

BAB IV

PENUTUP

4.1 Kesimpulan

Implementasi data warehouse dalam sektor perbankan merupakan langkah kritis untuk memungkinkan bank mengelola dan menganalisis volume besar data yang dihasilkan dari berbagai sumber, termasuk transaksi perbankan, interaksi pelanggan, dan data eksternal. Beberapa hal pentingnya seperti untuk kebutuhan analisis data, peningkatan layanan pelanggan, efisiensi operasional, keamanan, dan inovasi teknologi. dengan memanfaatkan data secara efektif, bank dapat menghasilkan wawasan yang bernilai tambah, meningkatkan kepuasan pelanggan, dan mencapai keunggulan kompetitif di pasar yang semakin kompetitif dalam menghadapi transformasi digital bank.

4.2 Saran

Disarankan untuk melakukan preprocessing data secara menyeluruh, seperti penanganan missing values, normalisasi, dan encoding variabel kategorikal agar prediksi lebih akurat. Selain itu, tuning parameter seperti kedalaman pohon dan kriteria pemisahan penting dilakukan untuk menghindari overfitting atau underfitting. Evaluasi model sebaiknya menggunakan metode seperti k-fold cross-validation agar hasil lebih stabil dan akurat. Jika data tidak seimbang, metode resampling atau penyesuaian class weights bisa diterapkan. Penggunaan fitur Feature Importance juga membantu dalam memahami variabel yang paling berpengaruh.

Namun, model Decision Tree cenderung overfitting, terutama jika pohon terlalu dalam. Penggunaan metode ensemble seperti Random Forest dapat menjadi solusi. Selain itu, Decision Tree kurang robust terhadap noise dan tidak efisien untuk dataset besar. Model juga sensitif terhadap perubahan kecil dalam data, sehingga metode yang lebih stabil sebaiknya dipertimbangkan jika data sering berubah.

Dengan memperhatikan hal-hal tersebut, performa model Decision Tree di RapidMiner dapat lebih optimal dan akurat.

DAFTAR PUSTAKA

- Arde, A. P., Marzal, J., & Saputra, E. (2021). EVALUASI PENERIMAAN PENGGUNAAN SISTEM INFORMASI AKADEMIK (SIKAD) UNIVERSITAS JAMBI MENGGUNAKAN UNIFIED THEORY OF ACCEPTANCE AND USE OF TECHNOLOGY (UTAUT). *Jurnal Rekayasa Informasi*, 10(1), 13–22.
- Melani, F., & Sulastri, S. (2023). Analisis Perbandingan Klasifikasi Algoritma CART dengan Algoritma C 4.5 Pada Kasus Penderita Kanker Payudara. *Jurnal Tekno Kompak*, 17(1), 171–183. <https://doi.org/10.33365/jtk.v17i1.2379>
- Pratasik, S. (2019). Perancangan Sistem Business Intelligence Pada Palang Merah Indonesia Daerah Sulawesi Utara. *Jurnal Frontiers*, 2(2), 199–209.
- Rofiani, R., Oktaviani, L., Vernanda, D., & Hendriawan, T. (2024). Penerapan Metode Klasifikasi Decision Tree dalam Prediksi Kanker Paru-Paru Menggunakan Algoritma C4.5. *Jurnal Tekno Kompak*, 18(1), 126–139. <https://doi.org/10.33365/jtk.v18i1.3525>
- Wahid, M. A. R., Nugroho, A., & Anshor, A. H. (2023). Prediksi Penyakit Kanker Paru-Paru Dengan Algoritma Regresi Linier. *Jurnal BIT*, 4(1), 63–74.
- Wirdasari, D., & Calam, A. (2011). Penerapan Data Mining Untuk Mengolah Data Penempatan Buku Di Perpustakaan Smk Ti Pab 7 Lubuk Pakam Dengan Metode Association Rule. *Jurnal Saintikom*, 10(2), 137–150.