

PENERAPAN ALGORITMA DECISION TREE UNTUK KREDIT SKORING DI BANK

Miftakhur Roziqi, Afifah Nurul Izzati

Teknik Informatika, Universitas Kahuripan Kediri
Jalan PB Sudirman No.25 Pare Kediri, Indonesia
miftakhurroziqi@students.kahuripan.ac.id

ABSTRAK

Dalam perbankan kredit adalah sumber utama penghasilan sekaligus merupakan risiko yang paling tinggi dalam operasional bank. Semakin ketatnya persaingan dalam dunia perbankan terutama dalam pemberian kredit, perbankan kerap kali menghiraukan prinsip kehati-hatian (*prudential bank*) terutama prinsip 5C (*character, capital, capacity, condition, dan collateral*). Permasalahan pada salah satu perbankan yaitu Kop. BPR Bumi Arta dimana dalam menentukan kredit scoring untuk pemberian kredit belum menerapkan prinsip kehati-hatian (*prudential bank*) dalam hal ini prinsip 5C, dari permasalahan tersebut maka diperlukan penerapan algoritma decision tree C4.5 untuk kredit scoring sehingga dapat memberikan prediksi yang lebih akurat tentang kemungkinan calon debitur gagal membayar kewajibannya serta mengurangi risiko kredit bermasalah dan kerugian financial. Metode penelitian yang digunakan adalah model *Cross-Standard Industry for Data Mining (CRISP-DM)*. Dari hasil penelitian ini didapatkan akurasi decision tree C4.5 sebesar 72,73% dari data sebanyak 36 record yang dibagi menjadi dua 70% data training dan 30% data testing.

Kata kunci : *Decision Tree, Algoritma C4.5, Koperasi, Bank Perkreditan Rakyat, Prinsip 5C*

1. PENDAHULUAN

Bank adalah badan usaha yang menghimpun dana dari masyarakat dalam bentuk simpanan dan menyalurkannya kepada masyarakat dalam bentuk kredit dan atau bentuk lainnya dalam rangka meningkatkan taraf hidup rakyat banyak [1].

Dalam perbankan kredit adalah sumber utama penghasilan, sekaligus merupakan risiko yang paling tinggi dalam kegiatan bank. Kredit bermasalah merupakan permasalahan yang cukup mendasar dalam perbankan, semakin besar kredit bermasalah maka akan berdampak pada dana operasional bank tersebut yang secara otomatis akan berdampak pada risiko likuiditas bank yang dapat menurunkan kepercayaan masyarakat terhadap kinerja bank tersebut.

Semakin ketatnya persaingan dalam dunia perbankan terutama dalam pemberian kredit, perbankan kerap kali menghiraukan prinsip kehati-hatian. Untuk memberikan kredit kepada calon debitur perbankan pada tahap awal akan melakukan kredit scoring terlebih dahulu terhadap calon debiturnya. Kredit scoring merupakan salah satu metode yang digunakan untuk menilai kelayakan dari calon debitur, pada umumnya bank akan mempersiapkan data-data dari calon debitur yang dapat menggambarkan kondisi calon debitur antara lain data pribadi, riwayat kredit, usia, status, masa kerja, jabatan dalam pekerjaan, dan data keuangan yang digunakan sebagai indikator penilaian. Dalam melakukan kredit scoring diperlukan tahapan sistematis dan analisis kredit yang mencakup prinsip kehati-hatian (*prudential banking*) atas dasar tersebut perbankan wajib berpedoman pada prinsip 5C (*character, capital, capacity, condition, dan collateral*) [2]. C yang pertama yaitu *Character* (Karakter) Karakter dalam hal ini adalah sifat atau watak calon debitur, untuk meyakinkan pihak bank bahwa sifat atau watak calon debitur benar-benar dapat

dipercaya. C yang kedua yaitu *Capacity* (Kapasitas) adalah bank mencoba melihat kemampuan calon debitur dalam mengembalikan kewajiban. C yang ketiga yaitu *Capital* (Permodalan) Tahap ini bank akan melihat seberapa besar modal yang dimiliki calon debitur dalam menjalankan usahanya. C yang keempat yaitu *Condition* (Kondisi Finansial) bank akan melihat kestabilan finansial dari calon debitur, yang bertujuan untuk memprediksi prospek usaha dimasa yang akan datang. C yang kelima yaitu *Collateral* (Jaminan) Berupa jaminan fisik ataupun non fisik yang diberikan calon debitur.

Pada penelitian ini mengangkat permasalahan pada salah satu perbankan yaitu Kop. BPR Bumi Arta dimana terdapat pemberian kredit yang tidak sesuai dengan prinsip kehati-hatian, dari masalah yang diteliti terdapat penentuan kredit scoring yang tidak sesuai yaitu belum menerapkan prinsip 5C (*character, capital, capacity, condition, dan collateral*)

Dari permasalahan diatas maka penerapan kredit scoring dengan menggunakan algoritma decision tree dapat dijadikan sebagai solusinya, sehingga dapat memberikan prediksi yang lebih akurat tentang kemungkinan calon debitur gagal membayar kewajibannya serta mengurangi risiko kredit bermasalah dan kerugian financial, penelitian kredit scoring juga sudah pernah dilakukan oleh peneliti sebelumnya namun belum menambahkan faktor 5C. Menurut Aryo Paryogo Nawary dan kawan-kawan dalam penelitiannya menyebutkan bahwa tingkat akurasi algoritma C4.5 sudah mencapai tingkat baik sehingga dapat meningkatkan proses ketelitian, klasifikasi dan prediksi [3], sedangkan menurut Dison Librado dan kawan-kawan menyebutkan bahwa algoritma C4.5 mampu menghasilkan tingkat akurasi 100% dibandingkan algoritma naive bayes [4]. Dari hasil penelitian sebelumnya yang menyatakan bahwa

algoritma decision tree menghasilkan performa yang baik, namun dari penelitian sebelumnya belum memperhitungkan prinsip-prinsip 5C tersebut.

2. TINJAUAN PUSTAKA

2.1. Data Mining

Adalah suatu proses ekstraksi pengetahuan atau menggali nilai tambah berupa informasi yang berharga dari suatu sumber data besar dan kompleks, sedangkan tujuan utamanya adalah mengidentifikasi pola, keterkaitan, atau informasi yang belum diketahui secara langsung dalam suatu data, sehingga dapat menambah wawasan yang lebih bernilai. Dalam penggunaannya data mining melibatkan berbagai bentuk teknik antara lain statistik, matematis, kecerdasan buatan untuk dapat menganalisis secara otomatis. Dari data mining dapat digunakan untuk mendukung pengambilan keputusan, tren pasar, meningkatkan efisiensi operasional dan merumuskan strategi bisnis[5].

2.2. Decision Tree

Decision Tree mempunyai konsep dasar merubah data menjadi suatu pohon keputusan dengan aturannya[6]. Atribut - atribut yang dipilih akan menghasilkan suatu partisi dengan data yang lebih seragam yang dapat menghasilkan sebuah pohon keputusan yang sederhana. Pohon keputusan terdiri dari beberapa kumpulan aturan yang tujuannya untuk membagi sejumlah populasi yang heterogen menjadi lebih kecil dan lebih homogen dengan memperhatikan variabel tujuannya. Dalam struktur pohon dimana setiap node pohon akan mempresentasikan hasil atribut yang telah dilakukan pengujian, kemudian setiap cabang merupakan pembagian dari hasil uji, dan node daun (leaf) akan mempresentasikan kelompok dalam kelas tertentu.

Level node teratas dari sebuah Decision Tree biasanya disebut sebagai akar (root) berupa atribut yang memiliki pengaruh dalam suatu kelas tertentu. Decision Tree melakukan pencarian secara top down untuk solusinya, proses pengklasifikasi data yang belum diketahui nilai atribut akan diuji dengan cara mencari jalur dari node akar (root) sampai dengan node akhir (daun) setelah itu akan dilakukan prediksi kelas yang telah dimiliki oleh suatu data tertentu. Secara sederhana Decision Tree dapat diartikan sebagai metode klasifikasi pada Text Mining[7].

2.3. Algoritma C45

Adalah salah satu algoritma pohon keputusan yang populer yang dikembangkan oleh Ross Quinlan pada tahun 1993 [8].

Algoritma ini adalah penerus dari algoritma ID3 (Iterative Dichotomiser 3) yang juga dikembangkan oleh Quinlan. Nama C4.5 merujuk pada versi keempat dari sistem pembelajaran klasifikasi yang dibangun oleh Quinlan. Kemudian hubungan antara C4.5 dengan Decision Tree adalah adalah salah satu implementasi dari algoritma Decision Tree, spesifiknya C4.5 adalah

implementasi yang menggunakan beberapa teknik khusus, seperti pemilihan fitur yang optimal dengan menggunakan metrik seperti Gain Ratio, penanganan data non-biner, penanganan nilai yang hilang, dan pruning pohon untuk mencegah overfitting atau bisa disebut juga sebagai salah satu contoh algoritma Decision Tree yang memperkuat pemahaman tentang bagaimana Decision Tree dapat diterapkan dan dikembangkan dengan teknik-teknik khusus untuk meningkatkan kinerjanya dalam pembelajaran mesin. Algoritma C4.5 untuk memperoleh nilai *entropy*, nilai *information gain*, nilai *split info*, dan nilai *gain ratio* dijabarkan sebagai berikut :

Menghitung nilai *entropy*

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

Keterangan :

- S = Himpunan kasus
- n = Jumlah partisi S
- p_i = proporsi S_i terhadap S

Menghitung nilai *information gain*

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

Keterangan :

- S = Himpunan kasus
- A = Fitur
- n = Jumlah partisi atribut A
- $|S_i|$ = Proporsi S_i terhadap S
- $|S|$ = Jumlah kasus dalam S

Menghitung nilai *Split info*

$$Split Info(S, A) = \sum_{i=1}^n \frac{|S_i|}{|S|} \log_2 \frac{|S_i|}{|S|}$$

Keterangan :

- S = ruang data sampel yang digunakan
- A = Atribut
- S_i = Jumlah sampel untuk atribut i

Menghitung nilai *Gain Ratio*

$$Gain Ratio(S, A) = \frac{Gain(S, A)}{Split Info(S, A)} \tag{4}$$

2.4. Rapid Miner

Adalah platform analisis data yang kuat dan serbaguna yang digunakan untuk pemodelan prediktif, analisis desain, penambangan data, dan pembuatan model. Rapid miner merupakan perangkat lunak open source yang populer dalam bidang ilmu data dan analisis bisnis.

2.5. Kredit Scoring

Merupakan salah satu metode yang digunakan untuk menilai kelayakan dari calon debitur, sehingga perbankan dapat mengetahui calon debitur tersebut layak atau tidak untuk diberikan kredit. Selain itu kredit scoring dapat diartikan sebagai proses penilaian risiko kredit bagi calon peminjam, sedangkan tujuan

kegiatan kredit scoring adalah untuk memprediksi kemungkinan pembayaran kembali pinjaman oleh debitur [9].

2.6. Kop. BPR Bumi Arta

Adalah salah satu penyedia jasa keuangan yang berpusat di Kabupaten Jombang, yaitu di Jl. Irian Jaya No. 25 Cukir Diwek Jombang, Kop. BPR Bumi Arta merupakan Bank yang melaksanakan kegiatan usahanya secara konvensional yang dalam kegiatannya tidak memberikan jasa dalam lalu lintas pembayaran. Kop. BPR Bumi Arta didirikan pada tahun 1990 dengan nomor ijin usaha 6782/BH/II/90 yang telah dikeluarkan oleh Kepala Kantor Wilayah Departemen Koperasi Jawa Timur. Selain diawasi Dinas Koperasi, Kop. BPR Bumi Arta telah terdaftar dan diawasi oleh Otoritas Jasa Keuangan dalam kegiatannya menghimpun dana dari masyarakat dan menyalurkannya dalam bentuk kredit dan penempatan pada bank lain

2.7. Otoritas Jasa Keuangan

Adalah lembaga independen yang mempunyai fungsi, tugas, dan wewenang pengaturan, pengawasan, pemeriksaan dan penyidikan. OJK dibentuk berdasarkan UU Nomor 21 Tahun 2011 yang berfungsi menyelenggarakan sistem pengaturan dan pengawasan yang terintegrasi terhadap keseluruhan kegiatan di dalam sektor jasa keuangan[10].

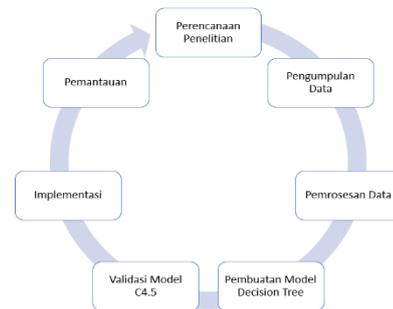
3. METODE PENELITIAN

Metode penelitian menggunakan model Cross-Standard Industry for Data Mining (CRISP-DM), CRISP-DM merupakan standart yang dikembangkan sejak tahun 1996 di Eropa, CRISP-DM bukan merupakan satu-satunya standart dalam data mining, namun merupakan model yang terpopuler saat ini [11], dalam CRISP-DM terdapat 6 fase yaitu:

- a. *Bussiness Understanding*
Mendefinisikan masalah dengan sederhana dan tepat atau bisa disebut sebagai fase pemahaman penelitian. Berdasarkan periode data desember 2023 kredit bermasalah lebih tinggi dari pada periode sebelumnya, hal ini sebagai acuan penelitian bahwa kurangnya analisa yang akurat dalam kredit scoring.
- b. *Data Understanding*
Menggabungkan data yang dimiliki dengan data yang akan diperlukan untuk penelitian, sehingga akan diperoleh data baik untuk penelitian.
- c. *Data Preparation*
Setelah *Data Understanding* kita dapatkan langkah berikutnya adalah *Data Perparation*, untuk menghasilkan data yang berkualitas terdapat langkah-langkah yang diperlukan dalam proses data preparation antara lain data *validation*, data *integration* dan *transformation*, data *size reduction* and *dicretization*.

- d. *Modeling*
Pada fase ini data training akan diproses untuk menentukan beberapa aturan yang akan membentuk sebuah pohon keputusan. Model yang akan digunakan yaitu algoritma decision tree C4.5.
- e. *Evaluation*
Evaluasi atau fase klasifikasi pada fase ini akan ditentukan pegujian untuk mendapatkan hasil yang akurat, yaitu dengan melihat hasil akurasi algoritma decision tree C4.5.
- f. *Deployment*
Fase terakhir adalah fase deployment yaitu menerapkan hasil algoritma decision tree C4.5 untuk menghasilkan pola kredit scoring.

3.1. Tahapan Penelitian



Gambar 1. Tahapan penelitian

Dari tahapan penelitian pada gambar 1, diuraikan sebagai berikut :

- a. *Perencanaan Penelitian*
Dalam perencanaan penelitian ini diperlukan identifikasi sumber daya yang diperlukan, data historis kredit, perangkat lunak, analisis data dan kemampuan teknis, untuk penelitian ini sumber daya yang diperlukan adalah bagian analisis kredit/kabag kredit pada Kop. BPR Bumi Arta, perangkat lunak yang diperlukan adalah Ms. Excel sebagai pengolahan data.
- b. *Pengumpulan Data*
Tahap ini peneliti mengumpulkan histori kredit yang mengandung informasi nominal pengajuan, jangka waktu, status perkawinan, pekerjaan, jumlah penghasilan, character, capacity, capital, condition, collateral, keterangan.
- c. *Pemrosesan Data*
Setelah data dikumpulkan maka data akan diproses untuk membersihkan nilai yang tidak sesuai misalnya tidak valid, kemudian mengkonversi variabel-variabel yang telah disiapkan untuk penelitian berdasarkan himpunan dan atribut, setelah itu membagi menjadi set pelatihan dan set pengujian.

Tabel 1. Penentuan value atribut

Atribut	Keterangan
Nominal Pengajuan	0-5 Jt (Kecil) >5 Jt <=25 Jt (Sedang) >25 Jt <=100 Jt (Besar) >100 Jt (Sangat Besar)
Jangka Waktu	1 – 6 Bln (Pendek) >6 – <=24 Bln (Cukup Lama) >24 – <=36 Bln (Lama) >36 Bln (Sangat Lama)
Status Perkawinan	Belum Kawin, Kawin, Cerai
Pekerjaan	PNS ,Petani ,TNI ,Wirausaha, Ibu Rumah Tangga, Lain-Lain
Jumlah Penghasilan	>= 2 Jt (Rendah) > 2 Jt s/d <=5 Jt (Sedang) > 5 Jt (Tinggi)
Character	Baik, Cukup Baik, Kurang Baik, Jelek
Capacity	Ada, Tidak
Capital	Ada, Tidak
Condition	Baik, Cukup, Kurang
Collateral	Mengkover, Tidak Mengkover
Keterangan	Layak, Tidak Layak

Dari tabel 1 dapat dilihat bahwa data yang dikumpulkan akan dikonversi sesuai dengan atribut masing-masing misalnya adalah atribut nominal pengajuan, jangka waktu, jumlah penghasilan dari data yang diterima atribut tersebut berupa angka sehingga perlu dikonversi sesuai dengan tabel 1, perangkat lunak yang digunakan untuk mengkonversi adalah Ms. Excel.

a. Pembuatan Model

Dengan menggunakan data yang telah disiapkan pembuatan model dapat ditentukan, dalam hal ini akan menggunakan model C4.5 untuk membuat struktur pohon keputusan dengan menggunakan software bantu rapid miner.

b. Validasi Model

Setelah model ditentukan maka langkah berikutnya adalah validasi model dengan menggunakan data pengujian, dari validasi model akan terlihat seberapa *accuracy*, *presisi*, dan *recall*.

c. Implementasi

Tahap selanjutnya adalah implementasi, implementasi akan dicoba menggunakan data yang sebenarnya untuk melakukan kredit scoring pada permohonan kredit yang masuk.

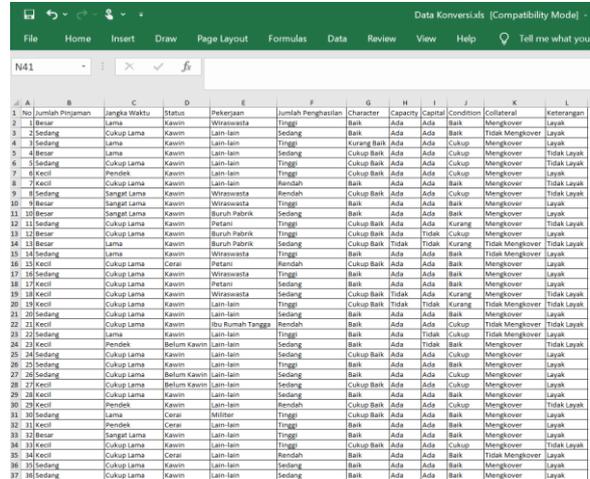
d. Pemantauan

Memantau seberapa banyak keptusan yang diambil oleh manajemen menggunakan algoritma decision tree ini.

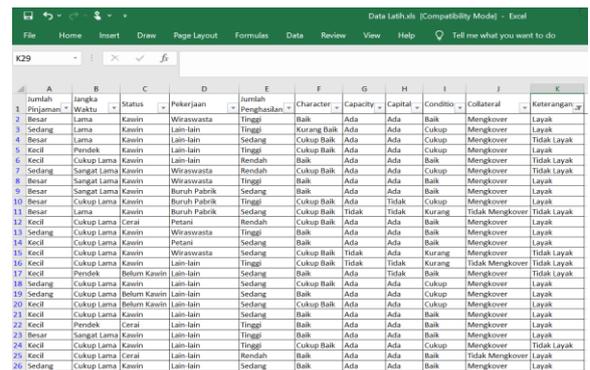
4. HASIL DAN PEMBAHASAN

Data yang telah terkumpul dan terkonversi sesuai dengan penjelasan pada tahapan penelitian sebanyak 38 record sesuai dengan tujuan penelitian maka akan dilakukan pemrosesan pada software bantu rapid miner, namun sebelum di proses terlebih dahulu dilakukannya teknik *preprocessing* yaitu:

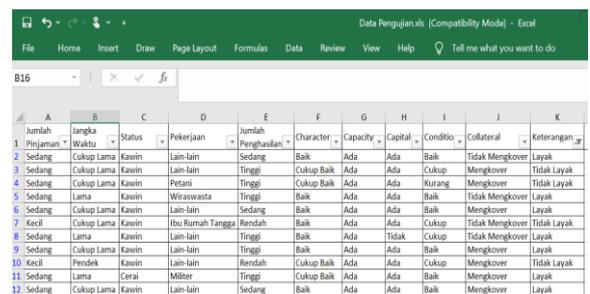
- Data *cleaning* yaitu membersihkan nilai yang kosong atau tidak konsisten.
- Data *integration* menyatukan tempat yang berbeda dalam satu data.
- Data *Reduction* menghapus atribut atau tupel yang tidak diperlukan[12].



Gambar 3. Data *cleaning*



Gambar 4. Data latihan

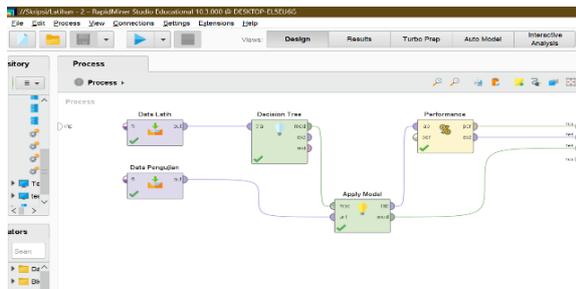


Gambar 5. Data pengujian

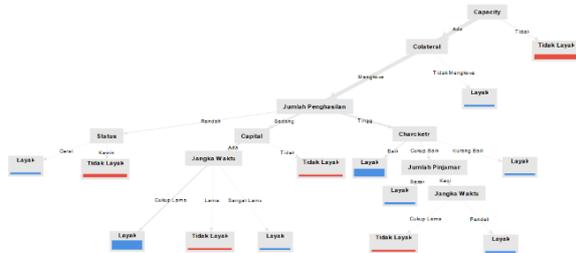
Gambar 3 menunjukkan bahwa data setelah dilakukan teknik *preprocessing* data *cleaning* yang semula data sebanyak 38 record menjadi 36 record, 2 record yang dihilangkan tersebut karena tidak memiliki value penghasilan perbulan dan keterangan, selanjutnya dari data *cleaning* tersebut akan dibagi menjadi dua yaitu 70% atau 25 record menjadi data latihan yang nampak pada gambar 4, dan 30% atau 11 record data pengujian pada gambar 5, dari data latihan tersebut akan ditentukan node terpilih sebagai node

awal, dengan pemikiran bahwa yang mempengaruhi kredit scoring dengan metode decision tree C4.5 adalah jumlah pinjaman, jangka waktu, status, pekerjaan, jumlah penghasilan, *character*, *capacity*, *capital*, *condition*, dan *collateral*, dengan demikian dari masing-masing atribut diatas dapat dihitung nilai *entropy*, *info gain*, *split info* dan *gain ratio*, *gain ratio* terbesar yang akan menjadi node awal.

4.1. Implementasi Rapid Miner



Gambar 6. Diagram proses rapid miner



Gambar 7. Pohon keputusan

Dari data latih yang telah diproses pada software rapid miner gambar 6 menghasilkan pohon keputusan yang nampak pada gambar 7, pohon keputusan tersebut menunjukkan bahwa node 1 dari data latih adalah *capacity*, node 2 *collateral*, node 3 jumlah penghasilan, node 4 status, node 5 *capital*, node 6 jangka waktu, node 7 *character*, node 8 jumlah pinjaman, node 9 jangka waktu. Dari pohon keputusan tersebut dan node-node yang dihasilkan peneliti memperoleh kesimpulan jika calon debitur memiliki *Capacity* (kemampuan bayar) maka akan dilakukan pengecekan terhadap *Collateral* (Jaminan) apakah *collateral* (jaminan) tersebut mengkover atau tidak jika tidak mengkover maka calon debitur akan mendapatkan kredit (dikarenakan calon debitur mempunyai kemampuan bayar dan pinjaman relatif kecil sehingga faktor jaminan tidak diperhitungkan), namun jika *collateral* (jaminan) mengkover maka akan memperhitungkan jumlah penghasilan jika penghasilan rendah maka akan memperhitungkan status perkawinan jika status cerai maka layak untuk diberikan kredit jika kawin maka tidak layak untuk diberikan pinjaman, kemudian kembali ke jumlah penghasilan jika jumlah penghasilan sedang maka akan memperhitungkan *Capital* (Modal) jika *Capital* (Modal) ada maka akan memperhitungkan jangka waktu jika jangka waktu cukup lama maka layak untuk

diberikan kredit, jika jangka waktu lama maka tidak layak untuk diberikan kredit, jika jangka waktu sangat lama maka layak untuk diberikan kredit, kembali ke *Capital* (Modal) jika *Capital* (Modal) tidak ada maka tidak layak untuk diberikan kredit, kemudian kembali lagi ke jumlah penghasilan jika jumlah penghasilan tinggi maka akan memperhitungkan *Character* (watak) calon debitur jika *Character* (watak) baik dan kurang baik maka layak untuk diberikan kredit, tapi jika *Character* (watak) cukup baik maka akan memperhitungkan jumlah pinjaman jika jumlah pinjaman besar maka layak untuk diberikan kredit, jika jumlah kecil maka akan memperhitungkan jangka waktu jika jangka waktu cukup lama maka tidak layak untuk diberikan pinjaman dan jika jangka waktu pendek maka layak untuk diberikan kredit, kemudian kembali ke *Capacity* (kemampuan bayar) jika *Capacity* (kemampuan bayar) tidak ada maka tidak layak untuk diberikan kredit.

Row No.	Keterangan	prediction(Keterangan)	confidence(Layak)	confidence(Tidak Layak)
1	Layak	Layak	1	0
2	Tidak Layak	Layak	0.667	0.333
3	Tidak Layak	Layak	0.667	0.333
4	Layak	Layak	1	0
5	Layak	Layak	1	0
6	Tidak Layak	Layak	1	0
7	Layak	Layak	1	0
8	Layak	Layak	1	0
9	Tidak Layak	Tidak Layak	0	1
10	Layak	Layak	0.667	0.333
11	Layak	Layak	1	0

Gambar 8. Hasil data pengujian

Dari gambar 8 menunjukkan bahwa pada data pengujian terdapat 3 record yang semula dengan kelas tidak layak, namun pada proses rapid miner 3 record tersebut diprediksi menjadi kelas layak dikarenakan confidence(layak) lebih besar dari pada confidence(tidak layak).

accuracy: 72.73%

	true Layak	true Tidak Layak	class precision
pred. Layak	7	3	70.00%
pred. Tidak Layak	0	1	100.00%
class recall	100.00%	25.00%	

Gambar 9. Performance

Performance pada data pengujian dapat dilihat pada gambar 9 dimana pada pengujian tersebut *accuracy* yang dihasilkan adalah sebesar 72.73%, *class recall true* layak 100%, *true* tidak layak 25%, sedangkan *class precision* prediksi layak adalah 70% tidak layak 100%.

4.2. Pembuktian

Gambar 10. Perhitungan node pertama

Dari perhitungan pada gambar 10 diketahui gain ratio yang terbesar adalah capacity 0.42304 sehingga atribut capacity menjadi node awal (root node), dengan menghasilkan keputusan jika capacity ada maka akan dilakukan ekstraksi pada atribut berikutnya, jika capacity tidak ada maka tidak layak, proses berikutnya adalah menentukan node 2 sebagai akar untuk proses perhitungan tidak ada yang berbeda dengan node 1, perlu dipahami bahwa capacity tidak perlu di hitung kembali dikarenakan sudah menjadi node awal, berikutnya adalah mengextraksi kelas ada pada atribut capacity dikarenakan masih memiliki entropy sebesar 0.773227, sedangkan kelas tidak, tidak akan dilanjutkan ekstraksi dikarenakan sudah tidak memiliki nilai entropy.

Gambar 11. Perhitungan node 2

Dari perhitungan pada gambar 11 dapat diketahui bahwa atribut collateral memiliki nilai gain ratio tertinggi sebesar 0.065083 sehingga atribut ini menjadi root node 2 dengan menghasilkan keputusan jika collateral mengcover maka akan dilakukan ekstraksi atribut berikutnya, jika collateral tidak mengcover maka tidak layak, kelas mengcover masih memiliki entropy sebesar 0.791858 dan akan dilakukan ekstraksi pada tahap berikutnya, sedangkan kelas tidak mengcover tidak lagi dilakukan ekstraksi.

Gambar 12. Perhitungan node 3

Pada gambar 12 diketahui bahwa gain ratio yang tertinggi adalah atribut jumlah penghasilan 0.081083 sehingga atribut ini menjadi root node 3, yang menghasilkan keputusan jika jumlah penghasilan rendah, sedang, tinggi akan dilakukan ekstraksi dikarenakan kelas-kelas tersebut memiliki nilai entropy perlu diketahui yang dilakukan ekstraksi berikutnya adalah kelas rendah dikarenakan memiliki nilai entropy tertinggi 0.918296, sedang 0.764205, tinggi 0.503258.

Gambar 13. Perhitungan node 4

Gambar 13 menunjukkan perhitungan node ke 4 yang dihasilkan dari ekstraksi pada node ke 3 atribut jumlah penghasilan kelas rendah, dari perhitungan tersebut yang memiliki gain ratio yang tertinggi adalah status perkawinan dengan gain ratio 1. Pada perhitungan node ke 4 ini menghasilkan keputusan jika status perkawinan kawin maka tidak layak, sedangkan jika status perkawinan cerai maka layak, untuk status perkawinan belum kawin tidak dipehitungkan dikarenakan tidak memiliki nilai layak atau tidak layak, status perkawinan ini tidak lagi di ekstraksi lagi dikarenakan semua kelas entropy menghasilkan nilai 0, sehingga pada node berikutnya adalah mengextraksi dari atribut jumlah penghasilan dengan kelas sedang.

Gambar 14. Perhitungan node 5

Dari gambar 14 menunjukkan dari hasil *extraksi atribut* jumlah penghasilan dengan kelas sedang, dari perhitungan diatas dapat diketahui bahwa *gain ratio* yang tertinggi adalah *atribut capital* dengan *gain ratio* 0.558433 yang menghasilkan keputusan jika *capital* ada layak sedangkan *capital* tidak maka tidak layak, kelas ada pada attribut *capital* akan diextraksi dikarenakan masih memiliki *entropy* 0.54356 untuk kelas tidak, tidak akan dilakukan *extraksi* dikarenakan sudah tidak memiliki *entropy*.

Gambar 15. Perhitungan node 6

Gambar 15 menunjukkan dari hasil *extraksi atribut capital* dengan kelas ada, dari perhitungan diatas dapat diketahui bahwa *gain ratio* yang tertinggi adalah *atribut* jangka waktu dengan *gain ratio* 0.528270 yang menghasilkan keputusan jika jangka waktu pendek,lama maka tidak layak sedangkan jika jangka waktu cukup lama dan sangat lama maka layak, attribut jangka waktu ini tidak akan diextraksi lagi pada *root node atribut* jumlah penghasilan sedang dikarenakan semua kelas tidak memiliki *entropy*. Berikutnya adalah mengextraksi *atribut* jumlah penghasilan dengan kelas tinggi.

Gambar 16. Perhitungan node 7

Dapat diketahui pada gambar 16 bahwa perhitungan *extraksi atribut* jumlah penghasilan kelas tinggi yang memiliki *gain ratio* yang tertinggi adalah *atribut character* dengan *gain ratio* 0.145867 yang menghasilkan keputusan jika *character* baik, kurang baik maka layak, jika *character* cukup baik maka akan dilakukan *extraksi* untuk menghasilkan *node* berikutnya.

Gambar 17. Perhitungan node 8

Selanjutnya pada gambar 17 yaitu hasil perhitungan dari *extraksi atribut character* dengan kelas cukup baik, dari perhitungan tersebut *gain ratio* adalah sama untuk itu diambil atribut yang paling depan yaitu *atribut* jumlah pinjaman dengan menghasilkan keputusan jika jumlah besar maka layak, jika jumlah kecil akan dilakukan *extraksi* untuk menghasilkan *node* berikutnya.

Gambar 18. Perhitungan node 9

Gambar 18 menunjukkan perhitungan hasil *extraksi* dari *atribut* jumlah pinjaman dengan kelas kecil, dari perhitungan tersebut *gain ratio* yang tertinggi adalah *atribut* jangka waktu sebesar 1, dengan menghasilkan keputusan jika jangka waktu pendek maka layak, jika jangka waktu cukup lama maka tidak layak sedangkan jangka waktu lama dan sangat lama tidak menghasilkan keputusan dikarenakan tidak terdapat nilai pada kelas tersebut. Pada node 9 ini juga merupakan perhitungan akhir dikarenakan semua *atribut gain ratio* adalah 0.

Dari hasil pembuktian ini dapat disimpulkan bahwa yang menjadi *node* awal adalah *capacity*, *node* kedua adalah *collateral node* ketiga adalah jumlah

penghasilan, *node* ke empat adalah status perkawinan, *node* kelima adalah *capital*, *node* keenam adalah jangka waktu *node* ketujuh adalah *character*, *node* kedelapan adalah jumlah pinjaman, *node* kesembilan jangka waktu telah sesuai dengan pohon keputusan yang dihasilkan dari software bantu rapid miner pada gambar 7.

Pada dasarnya kredit scoring dengan menerapkan prinsip 5C ini akan menguraikan banyak *node* akan tetapi hal tersebut telah sesuai dengan prinsip kehati-hatian (*prudential bank*).

5. KESIMPULAN DAN SARAN

Dari penelitian ini dapat disimpulkan bahwa algoritma *decision tree* C4.5 dapat diterapkan dalam penentuan kredit scoring dengan akurasi sebesar 72,73%. Sehingga dapat memberikan manfaat atau solusi kepada petugas perbankan atau pihak terkait dalam membuat penilaian terhadap analisa dalam pengajuan kredit dengan hasil yang lebih baik.

Saran dari penelitian ini adalah untuk penelitian selanjutnya adalah dapat menggunakan metode atau teknik yang lain sehingga didapatkan tingkat akurasi yang lebih baik lagi, serta perlunya pengembangan model algoritma sehingga jumlah kelas, kategori dan atribut dapat ditentukan secara optimal.

DAFTAR PUSTAKA

- [1] Indonesia. *Undang – Undang Nomor 4 Tahun 2023 Tentang Pengembangan Dan Penguatan Sektor Keuangan*. Jakarta
- [2] Otoritas Jasa Keuangan. 2024. *Peraturan Otoritas Jasa Keuangan No.1 Tahun 2024 Tentang Kualitas Aset Bank Perekonomian Rakyat*. Jakarta
- [3] Librado, D., Nasyuha, A., H., 2023. *Penerapan Data Mining Untuk Klasifikasi Penerima Kredit Dengan Perbandingan Algoritma Naïve Bayes dan Algoritma C4.5*, *Jurnal Media Informatika Budidarma* on (pp. 1952-1961).
- [4] Nawary, A.P., Kurniati, 2021. *Penerapan Data Mining Dalam Memprediksi Kelancaran Kredit Nasabah Menggunakan Algoritma C4.5*, *Bina Darma Conference on Computer Science* on (pp. 1041-1047).
- [5] Sarah, M.P., Hardinata, J.T., Kurniawan, H., Safii, M., Winajaya, R., 2022. *Implementasi Data Mining Menggunakan Algoritma Apriori Dalam Menentukan Persediaan Barang*. *Jurnal Sistem Informasi Teknologi Informasi dan Komputer* on (pp.51-61).
- [6] Kelvin., Mulyawan, B., Sutrisno, T., *Penerapan Algoritma C4.5 Untuk Penerimaan Kelayakan Kredit Pada Koperasi*. *Jurnal Ilmu Komputer dan Sistem Informasi* on (pp. 109-115).
- [7] Rusito, Firmansyah, M.T., 2016. *Implementasi Metode Decision Tree Dan Algoritma C4.5 Untuk Klasifikasi Data Nasabah Bank. INFOKAM* on (pp. 1-12).
- [8] Kasi, O.M., Muliantara, A., 2023. *Prediksi Paket Return Menggunakan Metode Decision Tree Menerapkan Algoritma C4.5 Berbasis Website, Jurnal Nasional Teknologi Informasi dan Aplikasinya* on (797-802).
- [9] Hakim, A.R., 2019. *Analisis Klasifikasi Credit Scoring Menggunakan Weighted Probabilistic Neural Network (WPNN)*. *Statistika* on (pp.63-70).
- [10] Oktariani, D., Ansyari, F., et al. 2023. *Booklet Perbankan Indonesia, Edisi 10. Tempat Publikasi : Otoritas Jasa Keuangan Departemen Perizinan dan Manajemen Krisis Perbankan, 2023, 12.*
- [11] Yusuf, D., Bahri, S., Larasati, A., 2021. *Decision Tree Menggunakan Algoritma C4.5 Untuk Analisa Kelayakan Pemberian Kredit*, *Jurnal Ilmiah Teknologi dan Informasi* on (pp. 97-106).
- [12] Rahayu, F., S., Samsiah, S., Hinggo, H., T., *Analisis Prinsip 5C dan 7P Dalam Pemberian Kredit Untuk Meminimalisir Kredit Bermasalah dan Meningkatkan Profitabilitas : Studi Kasus Pada Swamitra Pekan Baru, Seminar Nasional Ekonomi Bisnis dan Akuntansi, 20-27, Juli 14, 2021.*
- [13] Ramadhan, A.T., F.G.Hilmy, F., Puteri, N.R., Meirza, A., 2023. *Penerapan Algoritma Decision Tree Dalam Melakukan Analisis Klasifikasi Harga Handphone*, *Jurnal Sistem Informasi dan Ilmu Komputer* on (pp.195-206).
- [14] Iriadi, N., Nuraeni, N., 2016. *Kajian Penerapan Metode Klasifikasi Data Mining Algoritma C4.5 Untuk Prediksi Kelayakan Kredit Pada Bank Mayapada Jakarta*, *Jurnal Teknik Komputer AMIK BSI* on (pp. 132-137).