

PERBANDINGAN METODE ALGORITMA K-NN & METODE ALGORITMA C4.5 PADA ANALISA KREDIT MACET (STUDI KASUS PT TUNGMUNG TEXTILE BINTAN)

Ajeng Setianingrum^{1*}, Ayu Hindayanti², Dita Meilani Cahya³, Dini Silvi Purnia S.Kom,M.Kom⁴

Sistem Informasi, Universitas Nusa Mandiri

Indonesia

**Corresponding Autor: E-mail: dini.dlv@nusamandiri.ac.id*

ABSTRAK

Koperasi merupakan jasa peminjaman uang untuk membantu para anggotanya dalam memenuhi kebutuhan hidup sehari-hari. Pada proses simpan pinjam sering terjadi permasalahan yang muncul di dalam koperasi tersebut, contohnya para anggota terlambat untuk membayar cicilan karena sebab itu koperasi bisa menimbulkan permasalahan kredit macet. Pada penelitian ini ingin memprediksikan penyebab terjadinya kredit macet dengan memperhatikan data yang masuk dalam proses peminjaman uang. Metode yang akan digunakan untuk memprediksi kredit macet adalah perbandingan antara metode algoritma KNN dengan metode algoritma C4.5. Hasil akurasi algoritma C4.5 mendapat nilai yang lebih baik dibanding algoritma K-NN yaitu 61,64% sedangkan hasil akurasi K-NN yaitu 45,21%, dapat disimpulkan bahwa algoritma C4.5 lebih akurat untuk menentukan kredit macet.

Kata kunci : C4,5; K-NN; Koperasi; Kredit Macet; Perbandingan

ABSTRACT

Cooperative is a money lending service to help its members in meeting the needs of daily life. In the process of saving loans there are often problems that arise in the cooperative, for example, members are late to pay installments because therefore cooperatives can cause bad credit problems. In this study, we want to predict the cause of bad credit by paying attention to the data entered in the process of borrowing money. The method to be used to predict bad credit is a comparison between knn algorithm method and C4.5 algorithm method. The accuracy of the C4.5 algorithm scored better than the K-NN algorithm of 61.64% while the accuracy of K-NN was 45.21%, it can be concluded that the C4.5 algorithm is more accurate for determining bad credit.

Keywords: C4,5; K-NN; Cooperation; Bad Credit; C4.5; Compare

1. Latar Belakang

Menurut pasal 1 UU No. 25/1992, yang dimaksud dengan koperasi di Indonesia ialah badan usaha yang beranggotakan

seorang atau badan usaha hukum koperasi dengan melandaskan kegiatannya berdasarkan prinsip koperasi sekaligus

sebagai Gerakan ekonomi rakyat yang

Dalam menjalankan kegiatan simpan pinjam menimbulkan permasalahan yaitu kredit macet, karena kegiatan pemberian kredit ini berdasarkan kepercayaan terhadap yang menghimpun dana dari setiap anggotanya sebagai modal usaha.²

Dengan perkembangan teknologi yang semakin canggih membantu untuk mengembangkan kemampuan dalam mengelola data dan mengumpulkan data.

Salah satunya dengan menggunakan data mining. Data mining adalah serangkaian proses untuk menggali nilai tambah berupa informasi yang selama ini tidak diketahui secara manual dari suatu basis data yang sering disebut juga knowledge discovery database (KDD).³

Menurut (Tampubolon et al, 2013) Knowledge Discovery in Database (KDD) didefinisikan sebagai ekstraksi informasi potensial, implisit dan tidak dikenal dari sekumpulan data.⁴ Proses melibatkan hasil proses Data Mining (proses pengekstrak kecenderungan suatu pola data), kemudian mengubah hasilnya secara akurat menjadi informasi yang mudah dipahami.⁴

Menurut (Fayyad et al, 1996) kegunaan data mining dapat dibagi menjadi dua: deskriptif dan prediktif. Deskriptif berarti data mining digunakan untuk mencari pola pola yang dapat dipahami manusia yang menjelaskan karakteristik data.⁵

debitur.¹

Koperasi merupakan lembaga keuangan yang didirikan oleh orang perseorangan atau badan hukum koperasi,

Algoritma C4.5 adalah kelompok dari algoritma decision tree . algoritma ini mempunyai dua input yaitu *training samples* dan *testing samples*.⁶ Algoritma C4.5 akan menghasilkan pohon keputusan. Pohon keputusan atau decision tree adalah pohon yang digunakan sebagai prosedur penalaran untuk mendapatkan jawaban dari masalah yang dimasukkan.⁷ Algoritma KNN adalah menentukan kelas berdasarkan tingkat kesamaan yang tinggi dari data latih.⁸

Dalam penelitian ini menggunakan algoritma C4.5 karena area pengambilan keputusan yang kompleks dapat diubah menjadi lebih sederhana, dan dapat menghilangkan perhitungan yang tidak perlu karena proses pengujian hanya berdasarkan kriteria yang diperlukan saja.

Untuk algoritma KNN itu sendiri karena algoritma KNN mempunyai kelebihan, yaitu mempunyai akurasi yang tinggi, insensitive terhadap outlier, dan tidak ada asumsi terhadap data. selain itu algoritma KNN juga memiliki kekurangan, yaitu perlu memutuskan nilai K optimal, komputasi yang mahal, dan membutuhkan banyak memori.

Oleh sebab itu, karena banyak dari penelitian sebelumnya mendapatkan masalah yang sama yang membedakan hanya penggunaan metodenya saja, maka dari itu kami menggunakan metode algoritma C4.5 dan algoritma KNN sebagai bahan informasi dan referensi yang baru.

Pada penelitian yang telah dilakukan oleh (Tedi Permana, Amril Mutoi Siregar, Anis Fitri Nur Masruriyah, Ayu Ratna Juwita, 2020) dengan memprediksi hasil perbandingan kredit macet dengan algoritma C5.0 dengan hasil 86,67% dan

algoritma KNN dengan hasil 83,33%.⁹

Pada penelitian (Femi Dwi Astuti dan Mohammad Guntara, 2018) untuk klasifikasi data penduduk miskin dengan algoritma C4.5 94,71% dan algoritma KNN K=1 mempunyai performa yang lebih baik daripada K=10,100,1000.¹⁰

Pada penelitian (Yosep Nuryaman, 2018) untuk klasifikasi menentukan customer menggunakan algoritma C4.5 dengan hasil 84% dan KNN memberikan hasil 84,46% .¹¹

Berikut beberapa penelitian terdahulu yang berhubungan dengan penelitian skripsi ini sebagai bahan acuan dalam pengambilan metode.

Tabel 1.1 Penelitian Terdahulu

JUDUL	METODE ANALISIS	HASIL ANALISIS
Perbandingan algoritma C4.5, KNN, dan <i>Naive Bayes</i> untuk penentuan model klasifikasi penanggung jawab BSI <i>Entrepreneur Center</i> . ¹²	C4.5, KNN, dan <i>Naive Bayes</i>	Karyawan yang layak menjadi penanggung jawab BSI <i>Entrepreneur Center</i> pada masing-masing kampus Universitas BSI dengan metode <i>Naive Bayes</i> yang memiliki akurasi tinggi 80%
Perbandingan Algoritma <i>Naive Bayes</i> dan C4.5 Dalam Klasifikasi Data	<i>Naive Bayes</i> dan C4.5	Kasus penerimaan "Kartu Indonesia Sehat" kedua algoritma tersebut sama-sama

<p>Mining.¹³</p>		<p>efektif digunakan, lalu kasus penentuan pengajuan kartu kredit di sebuah bank C4.5 lebih efektif dibanding <i>Naive Bayes</i>. Kasus penerimaan usia kelahiran hasilnya <i>Naive Bayes</i> lebih baik dibanding C4.5. Sedangkan untuk kasus penentuan kelayakan calon anggota kredit pada koperasi <i>Naive Bayes</i> memberikan hasil yang baik untuk <i>precision</i>, tetapi untuk <i>recall</i> dan <i>accuracy</i> C4.5 memberikan hasil yang baik.</p>
<p>Perbandingan Metode <i>Naive Bayes Classifier</i> Dengan Metode <i>Decision Tree</i> (C4.5) Untuk Menganalisa Kelancaran Pembiayaan (study kasus KSPPS/ BMT AL-FADHILA).¹⁴</p>	<p><i>Naive Bayes Classifier</i> dan Metode <i>Decision Tree</i> (C4.5)</p>	<p>Metode <i>decision tree</i> (C4.5) memiliki tingkat akurasi lebih tinggi dan lebih efisien waktu yang lebih cepat daripada metode <i>Naive bayes classifier</i></p>
<p>Perbandingan Hasil Prediksi Kredit Macet Pada Koperasi Menggunakan Algoritma</p>	<p>algoritma <i>K-Nearest Neighbor</i> dan algoritma C5.0</p>	<p>hasil akurasi dari algoritma C5.0 mendapatkan nilai yaitu 86,67%. Sedangkan</p>

KNN dan C5.0. ⁹		untuk algoritma KNN mendapatkan nilai akurasi 83,33%, sehingga algoritma C5.0 lebih efektif untuk memprediksi kredit macet
----------------------------	--	--

Pada Koperasi PT. Tungmung Textile Bintan adanya pemberian pinjaman pada anggotanya yang tidak jarang menimbulkan permasalahan – permasalahan terutama masalah terlambatnya pembayaran angsuran yang sering juga disebut dengan tunggakan atau kredit macet. Pada penelitian yang dilakukan oleh tedi permana dengan memprediksi hasil perbandingan kredit macet di koperasi dengan menggunakan perbandingan algoritma KNN dan algoritma C5.0, penelitian tersebut mendapatkan hasil bahwa algoritma C5.0 lebih efektif untuk menentukan kredit macet daripada

algoritma KNN. Penelitian yang kami lakukan berfokus pada Perbandingan suatu metode dengan metode lainnya, oleh karena itu dalam penelitian ini intens untuk membahas perbandingan antara kedua metode, dan tujuan dari penelitian ini adalah untuk mengetahui metode manakah yang lebih baik untuk menentukan dan memecahkan masalah tentang kredit macet. maka pada penelitian ini akan mengambil judul “Perbandingan Metode Algoritma KNN & Metode Algoritma C4.5 Pada Analisa Kredit Macet (Studi Kasus PT TungMung Textile Bintan).

2. Metode Penelitian

Metode Penelitian berasal dari dua suku kata yaitu metode berasal dari Bahasa Yunani *methodos* yang berarti cara atau jalan yang ditempuh, dan penelitian berasal dari kata *research* “re” adalah kembali “search” mencari.¹⁵

Metode penelitian yang dilakukan dengan mengumpulkan data, antara lain :

1. Studi pustaka yang dilakukan untuk melengkapi penelitian ini yaitu dengan cara mengumpulkan referensi dari jurnal, buku-buku, internet sesuai dengan isi dari penyusunan skripsi ini
2. Wawancara, penulis melakukannya dengan cara tanya jawab seputar kredit macet dengan narasumber

atau bagian yang terkait di dalam koperasi PT. Tungmung Textile Bintang.

2.1 Identifikasi Masalah

Dalam penelitian ini terdapat masalah yang ada pada PT. Tungmung Textile Bintang yaitu terdapat banyaknya nasabah yang pembayarannya tidak sesuai dengan waktu yang ditetapkan dan mengganggu proses berjalannya koperasi tersebut.

2.2. Seleksi Data

Tahap ini dilakukan sebagai langkah awal dari suatu penelitian, penulis menyeleksi data-data pinjaman nasabah yang bermasalah pada koperasi PT. Tungmung Textile Bintang yaitu dari dana *emergency* dan dana siaga yang memiliki 73 data.

2.3. Pemilihan Metode Data Mining

Setelah menyeleksi data, penulis memilih metode data mining dengan menggunakan perbandingan antara metode algoritma KNN dan metode algoritma C4.5 untuk mengetahui algoritma mana yang lebih efektif untuk memprediksi kredit

macet untuk masalah tersebut.

2.4. Implementasi Algoritma KNN dan Algoritma C4.5

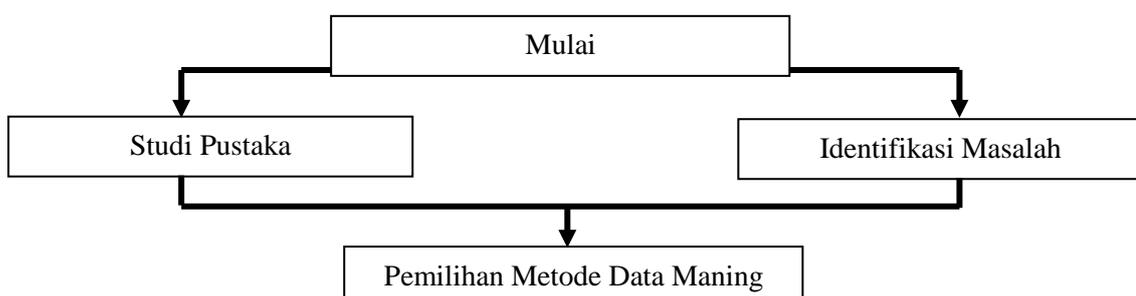
Setelah menentukan ingin menggunakan metode algoritma KNN dan metode algoritma C4.5, lalu penulis mengimplementasikan dengan menggunakan perhitungan manual untuk memudahkan menyelesaikan masalah kredit macet.

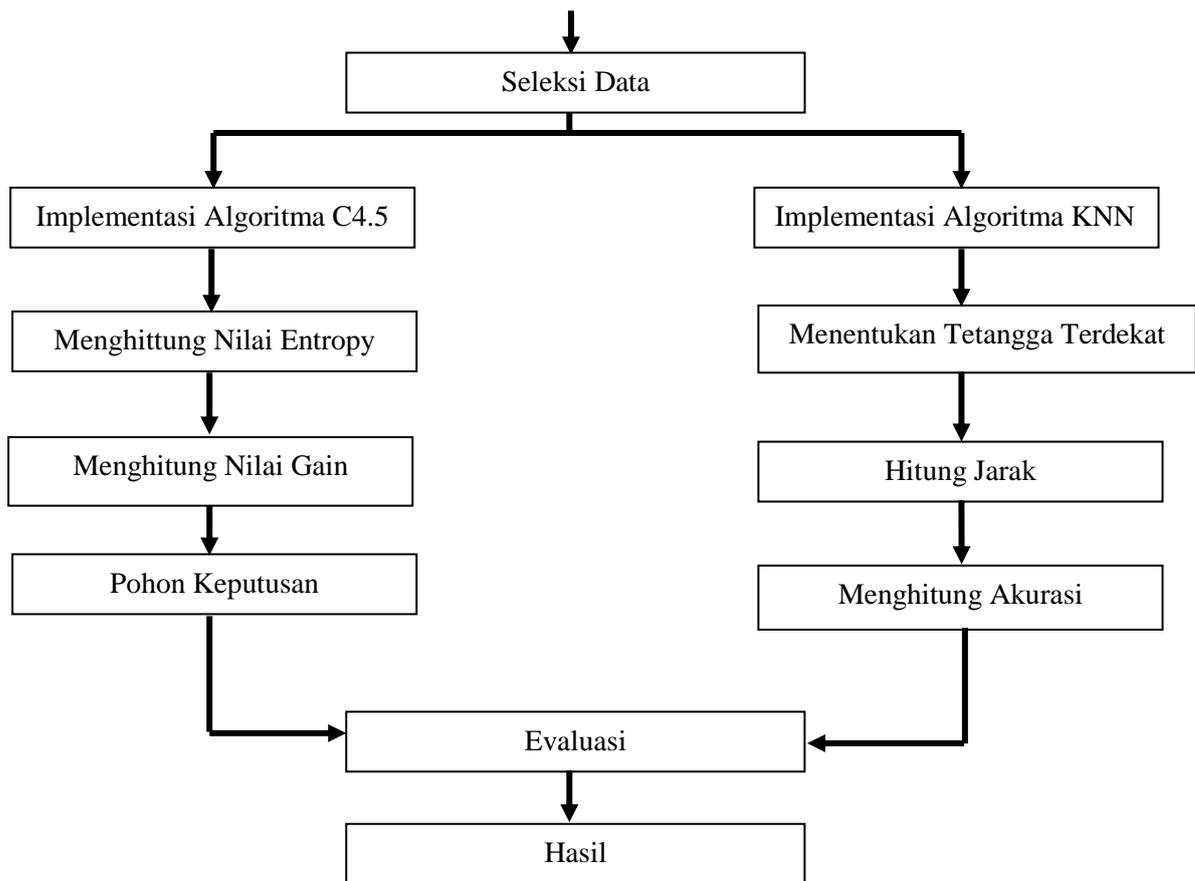
2.5. Validasi dan Evaluasi

Tahap ini dilakukan validasi dan evaluasi terhadap metode algoritma yang telah ditetapkan untuk mengetahui algoritma manakah yang memiliki tingkat akurasi lebih besar untuk menyelesaikan masalah tersebut.

2.6. Tahapan Penelitian

Alur Penelitian dimulai dengan identifikasi masalah, seleksi data, pemilihan metode data mining, implementasi algoritma C4.5 dan KNN , evaluasi lalu hasil.





Gambar 2.6. Alur Penelitian
(Penelitian 2021)

3. Hasil dan Pembahasan

3.1. Algoritma KNN

Dalam perhitungan manual KNN data latih yang diambil *datasets* yaitu nomor unit 1

sampai 10, sedangkan data yang diambil untuk data uji yaitu *datasets* nomor urut 11

Rumus perhitungan KNN adalah :

$$d(x,y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

Tabel 4.5. Gabungan Data Latih dan Data Uji

(Penelitian 2021)

NAMA	Jenis Kelamin	Jenis Pinjaman	Jumlah Pinjaman	Lama Cicilan	Keterangan
YENI HANDAYANI	2	1	1	3	TIDAK LANCAR
IDA FARIDA	2	1	2	1	TIDAK LANCAR
NANDANG SADIKIN	1	1	1	1	TIDAK LANCAR
SITI MARYANI	2	1	3	2	TIDAK LANCAR
NINING YUNINGSIH	2	1	1	3	LANCAR
ANDI BITIKDO PARSAORAN S	1	1	4	4	TIDAK LANCAR
RATIKAH	2	1	1	3	TIDAK LANCAR
ROHAYATI	2	1	1	1	TIDAK LANCAR
HAYATI SEPTIANAWTI	2	1	2	3	LANCAR
...
HENDRIK	1	1	2	3	?

Berikut Cara Perhitungan Eulidean Distance :

Tabel 4.6. Perhitungan *Eulidean Distance*

(Penelitian 2021)

Jenis Kelamin	Jenis Pinjaman	Jumlah Pinjaman	Lama Cicilan	Perhitungan Eulidean Distance
2	1	1	3	$\sqrt{((2-1)^2 + (1-1)^2 + (1-2)^2 + (3-3)^2)} = 1,414$
2	1	2	1	$\sqrt{((2-1)^2 + (1-1)^2 + (2-2)^2 + (3-3)^2)} = 2,236$
1	1	1	1	$\sqrt{((1-1)^2 + (1-1)^2 + (1-2)^2 + (1-3)^2)} = 2,236$
2	1	3	2	$\sqrt{((2-1)^2 + (1-1)^2 + (3-2)^2 + (2-3)^2)} = 1,732$
2	1	1	3	$\sqrt{((2-1)^2 + (1-1)^2 + (1-2)^2 + (3-3)^2)} = 1,414$
1	1	4	4	$\sqrt{((1-1)^2 + (1-1)^2 + (4-2)^2 + (4-3)^2)} = 2,236$
2	1	1	3	$\sqrt{((2-1)^2 + (1-1)^2 + (1-2)^2 + (3-3)^2)} = 1,414$
2	1	1	1	$\sqrt{((2-1)^2 + (1-1)^2 + (1-2)^2 + (1-3)^2)} = 2,449$

2	1	2	3	$\sqrt{((2-1)^2 + (1-1)^2 + (2-2)^2 + (3-3)^2)} = 1$
---	---	---	---	--

Parameter “K” yang digunakan yaitu 2, setelah menentukan Eulidean Distance tahap selanjutnya pada tabel 4.7. adalah menentukan kategori tetangga terdekat :

Tabel 4.7. Kategori Tetangga Terdekat
 (Penelitian 2021)

Jenis Kelamin	Jenis Pinjaman	Jumlah Pinjaman	Lama Cicilan	Eulidean Distance	Urutan Jarak	Apakah Termasuk 2-NN ?	Kategori Tidak untuk KNN
2	1	1	3	1,414	2	Ya (K<2)	Tidak Lancar
2	1	2	1	2,236	4	Tidak (K>2)	-
1	1	1	1	2,236	4	Tidak (K>2)	-
2	1	3	2	1,732	3	Ya (K<2)	Tidak Lancar
2	1	1	3	1,414	2	Ya (K<2)	Lancar
1	1	4	4	2,236	4	Tidak (K>2)	-
2	1	1	3	1,414	2	Ya (K<2)	Tidak Lancar
2	1	1	1	2,449	5	Tidak (K>2)	-
2	1	2	3	1	1	Ya (K<2)	Lancar

Setelah menentukan kategori tetangga terdekat tahap selanjutnya adalah mengetahui hasil prediksi KNN :

Tabel 4.8. Hasil Prediksi K-NN
 (Penelitian 2021)

Jenis Kelamin	Jenis Pinjaman	Jumlah Pinjaman	Lama Cicilan	Kategori Tidak untuk KNN
2	1	1	3	Tidak Lancar
2	1	2	1	-
1	1	1	1	-

2	1	3	2	Tidak Lancar
2	1	1	3	Lancar
1	1	4	4	-
2	1	1	3	Tidak Lancar
2	1	1	1	-
2	1	2	3	Lancar
1	1	2	3	Tidak Lancar

Accuracy :

accuracy: 45.00% +/- 16.94% (mikro: 45.21%)			
	true TIDAK LANCAR	true LANCAR	class precision
pred. TIDAK LANCAR	24	31	43.64%
pred. LANCAR	9	9	50.00%
class recall	72.73%	22.50%	

Gambar 3.1. Hasil Accuracy K-NN Rapid Miner

(Penelitian 2021)

3.2. Algoritma C4.5

Pada perhitungan Algoritma C4.5 ini menggunakan 73 data, dengan 4 atribut yaitu jenis kelamin, jenis pinjaman, jumlah pinjaman, dan lama cicilan

Untuk menghitung nilai entropy dengan rumus :

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

Untuk menghitung nilai gain dengan

rumus :

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

Hitunglah semua entropy dan gain sampai dengan selesai dan kemudian dapatkan node 1. Hasil perhitungan dapat dilihat seperti tabel dibawah ini :

Tabel 4.11 Perhitungan Gain Node 1

(Penelitian 2021)

ATRIBUT	JUMLAH	LANCAR	TIDAK	ENTROPY	GAIN
---------	--------	--------	-------	---------	------

		KASUS	Y	LANCAR		
				N		
Total		73	40	33	0.9934	
Jenis Kelamin	1	10	5	5	1.0000	0.0010
	2	63	35	28	0.9911	
Jenis Pinjaman	1	29	12	17	0.9784	0.0346
	2	44	28	16	0.9457	
Jumlah Pinjaman	1	11	4	7	0.9457	0.0299
	2	6	2	4	0.9183	
	3	10	5	5	1.0000	
	4	46	28	18	0.9656	
Lama Cicilan	1	9	3	6	0.9183	0.0668
	2	6	1	5	0.6500	
	3	12	7	5	0.9799	
	4	11	7	4	0.9457	
	5	35	22	13	0.9518	

Untuk *Gain Node* 1 berada pada atribut lama cicilan dengan *Gain* terbesar yaitu 0.0668 dan *Entropy* terkecil berada di kategori ke 2 dengan nilai 0.6500 yang akan digunakan untuk perhitungan *Gain Node* selanjutnya

Untuk mencari nilai entropy metode Algoritma C4.5 diatas sebagai berikut :

$$1. \text{ Entropy Total } (S) =$$

$$\left(\left(-\frac{40}{73} \right) \times \log_2 \left(\frac{40}{73} \right) \right) + \left(\left(-\frac{33}{73} \right) \times \log_2 \left(\frac{33}{73} \right) \right) = 0.9934$$

$$2. \text{ Entropy Jenis Kelamin 1 } (S) =$$

$$\left(\left(-\frac{5}{10} \right) \times \log_2 \left(\frac{5}{10} \right) \right) + \left(\left(-\frac{5}{10} \right) \times \log_2 \left(\frac{5}{10} \right) \right) = 1.000$$

$$3. \text{ Entropy Jenis Kelamin 2 } (S) =$$

$$\left(\left(-\frac{35}{63} \right) \times \log_2 \left(\frac{35}{63} \right) \right) + \left(\left(-\frac{28}{63} \right) \times \log_2 \left(\frac{28}{63} \right) \right) = 0.9911$$

$$4. \text{ Entropy Jenis Pinjaman 1 } (S) =$$

$$\left(\left(-\frac{12}{29} \right) \times \log_2 \left(\frac{12}{29} \right) \right) + \left(\left(-\frac{17}{29} \right) \times \log_2 \left(\frac{17}{29} \right) \right) = 0.9784$$

$$5. \text{ Entropy Jenis Pinjaman 2 } (S) =$$

$$\left(\left(-\frac{28}{44} \right) \times \log_2 \left(\frac{28}{44} \right) \right) + \left(\left(-\frac{16}{44} \right) \times \log_2 \left(\frac{16}{44} \right) \right) = 0.9457$$

$$6. \text{ Entropy Jumlah Pinjaman 1 } (S) =$$

$$\left(\left(-\frac{4}{11} \right) \times \log_2 \left(\frac{4}{11} \right) \right) + \left(\left(-\frac{7}{11} \right) \times \log_2 \left(\frac{7}{11} \right) \right) = 0.9457$$

$$7. \text{ Entropy Jumlah Pinjaman 2 } (S) =$$

$$\left(\left(-\frac{2}{6} \right) \times \log_2 \left(\frac{2}{6} \right) \right) + \left(\left(-\frac{4}{6} \right) \times \log_2 \left(\frac{4}{6} \right) \right) = 0.9183$$

$$8. \text{Entropy Jumlah Pinjaman 3 (S)} = \left(\left(-\frac{5}{10} \right) \times \log_2 \left(\frac{5}{10} \right) \right) + \left(\left(-\frac{5}{10} \right) \times \log_2 \left(\frac{5}{10} \right) \right) = 1.000$$

$$9. \text{Entropy Jumlah Pinjaman 4 (S)} = \left(\left(-\frac{28}{46} \right) \times \log_2 \left(\frac{28}{46} \right) \right) + \left(\left(-\frac{18}{46} \right) \times \log_2 \left(\frac{18}{46} \right) \right) = 0.9656$$

$$10. \text{Entropy Lama Cicilan 1 (S)} = \left(\left(-\frac{3}{9} \right) \times \log_2 \left(\frac{3}{9} \right) \right) + \left(\left(-\frac{6}{9} \right) \times \log_2 \left(\frac{6}{9} \right) \right) = 0.9183$$

$$11. \text{Entropy Lama Cicilan 2 (S)} = \left(\left(-\frac{1}{6} \right) \times \log_2 \left(\frac{1}{6} \right) \right) + \left(\left(-\frac{5}{6} \right) \times \log_2 \left(\frac{5}{6} \right) \right) = 0.6500$$

$$12. \text{Entropy Lama Cicilan 3 (S)} = \left(\left(-\frac{7}{12} \right) \times \log_2 \left(\frac{7}{12} \right) \right) + \left(\left(-\frac{5}{12} \right) \times \log_2 \left(\frac{5}{12} \right) \right) = 0.9799$$

$$13. \text{Entropy Lama Cicilan 4 (S)} = \left(\left(-\frac{7}{11} \right) \times \log_2 \left(\frac{7}{11} \right) \right) + \left(\left(-\frac{4}{11} \right) \times \log_2 \left(\frac{4}{11} \right) \right) = 0.9457$$

$$14. \text{Entropy Lama Cicilan 5 (S)} = \left(\left(-\frac{22}{35} \right) \times \log_2 \left(\frac{22}{35} \right) \right) + \left(\left(-\frac{3}{13} \right) \times \log_2 \left(\frac{3}{13} \right) \right) = 0.9518$$

Untuk mencari nilai *Gain* metode Algoritma C4.5 diatas sebagai berikut :

$$1. \text{Gain Jenis Kelamin (S,A)} = 0,9934 - \left(\left(\frac{10}{73} \right) \times 1.000 \right) + \left(\left(\frac{63}{73} \right) \times 0.9911 \right) = 0.0010$$

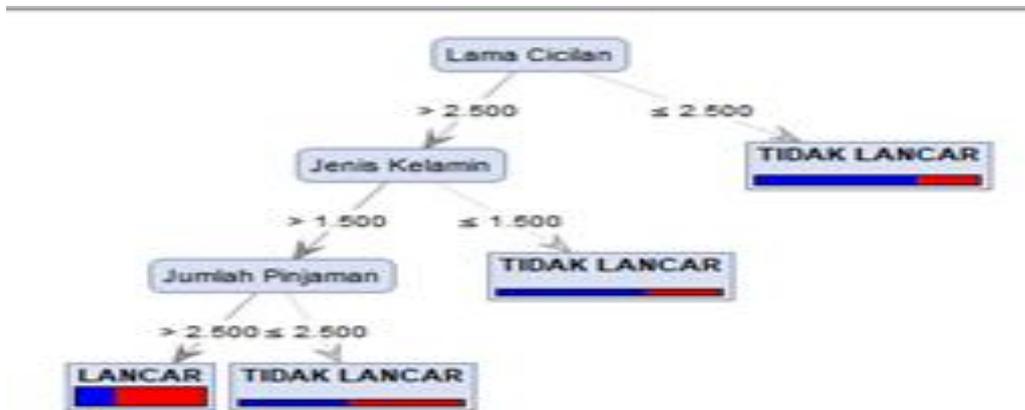
$$2. \text{Gain Jenis Pinjaman (S,A)} = 0,9934 - \left(\left(\frac{29}{73} \right) \times 0.9784 \right) + \left(\left(\frac{44}{73} \right) \times 0.9457 \right) = 0.0346$$

$$3. \text{Gain Jumlah Pinjaman (S,A)} = 0,9934 - \left(\left(\frac{11}{73} \right) \times 0.9457 \right) + \left(\left(\frac{6}{73} \right) \times 0.9183 \right) + \left(\left(\frac{10}{73} \right) \times 1.000 \right) + \left(\left(\frac{46}{73} \right) \times 0.9656 \right) = 0.0299$$

$$4. \text{Gain Lama Cicilan (S,A)} = 0,9934 - \left(\left(\frac{9}{73} \right) \times 0.9813 \right) + \left(\left(\frac{6}{73} \right) \times 0.6500 \right) + \left(\left(\frac{12}{73} \right) \times 0.9799 \right) + \left(\left(\frac{11}{73} \right) \times 0.9457 \right) + \left(\left(\frac{35}{73} \right) \times 0.9518 \right) = 0.0668$$

Perhitungan dilakukan terus menerus sehingga mendapatkan nilai *Gain Node* terbesar dan *Entropy* terkecil hingga selesai.

Pohon Keputusan :



Gambar 3.2. Hasil Pohon Keputusan *Rapid Miner*
 (Penelitian 2021)

Pembentukan pohon keputusan pada gambar 3.2 dilakukan dengan menghitung entropy masing-masing atribut, kemudian mencari information Gain Node terbesar dan yang terbesar itu digunakan sebagai simpul. Kemudian dilakukan proses berulang-ulang sampai pohon berhenti terbentuk.

Accuracy :

accuracy: 61.43% +/- 8.97% (mikro: 61.64%)			
	true TIDAK LANCAR	true LANCAR	class precision
pred. TIDAK LANCAR	15	10	60.00%
pred. LANCAR	18	30	62.50%
class recall	45.45%	75.00%	

Gambar 3.3. Hasil *Accuracy C4.5 Rapid Miner*
 (Penelitian 2021)

3.3. Kesimpulan

Berdasarkan hasil penelitian dengan judul “Perbandingan Metode Algoritma KNN dan Metode Algoritma C4.5 pada Analisa Kredit Macet(Studi Kasus Koperasi

PT. Tunggung Textile Bintan) dengan menggunakan 73 dataset dan 5 atribut. Algoritma KNN mendapatkan nilai akurasi sebanyak 45,21% sedangkan Algoritma C4.5 mendapatkan nilai akurasi sebanyak 61,64%. Penulis dapat menyimpulkan

bahwa Algoritma C4.5 lebih akurat untuk menentukan kredit macet daripada Algoritma KNN.

3.4. Saran

Berdasarkan konklusi maka penulis bisa memberikan saran menjadi:

1. Di penelitian selanjutnya bisa mencoba menggunakan dataset yang tidak sama dengan menggunakan jumlah yang besar sebagai akibatnya nilai data simpan pinjam koperasi yang dihasilkan akan mendapatkan nilai akurasi yang lebih tinggi.
2. Di penelitian selanjutnya dapat mencoba menggunakan teknik, algoritma serta *software* data mining yang berbeda sebagai akibatnya akan menerima berita yang pariatif.
3. Selain melakukan penelitian secara teoritis serta aplikatif, dibutuhkan pada penelitian berikutnya bisa mencoba untuk membuat suatu *software*.

3.5 References

1. Proaksi J. Analisis Faktor Internal Yang Mempengaruhi Kredit Macet. *J Proaksi*. 2020;2(2):128-134.
2. Nawangsari Y, Putra IS. Analisis Sistem Pengendalian Intern Pemberian Kredit Dalam Menurunkan Tingkat Kredit Macet Pada Koperasi Simpan Pinjam Kharisma Mitra Karya. *J Kompil*

- Ilmu Ekon*. 2016;8(1):67-80. <http://journal.stieken.ac.id/index.php/kompilek/article/view/278>
3. Vulandari RT. *Data Mining Teori Dan Aplikasi Rapidminer.*; 2017.
4. Sepri D, Afdal M. Analisa Dan Perbandingan Metode Algoritma Apriori Dan Fp-Growth Untuk Mencari Pola Daerah Strategis. *J Sist Inf Kaputama*. 2017;1(1):47-55.
5. Suryanto. *Data Mining Untuk Klasifikasi Dan Klasterisasi Data.*; 2017.
6. Atma YD, Setyanto A. Perbandingan algoritma c4.5 dan k-nn dalam identifikasi mahasiswa berpotensi drop out. *Metik J*. 2018;2(2):31-37.
7. Hariyanto M, Kholiq M, Yani A, Narti. Inti nusa mandiri. *Inti Nusa Mandiri*. 2020;14(2):133-138.
8. Lailil Muflikhah, Dian Eka Ratnawati RRMP. *Data Mining.*; 2018.
9. Permana T, Siregar AM, Masruriyah AFN, Juwita AR. Perbandingan Hasil Prediksi Kredit Macet Pada Koperasi. *Conf Innov Appl Sci Technol*. 2020;3(1):737-746. <http://publishing-widyagama.ac.id/ejournal-v2/index.php/ciastech/article/view/1970>
10. Astuti FD, Guntara M. Analisis Performa Algoritma K-NN Dan C4.5

-
- Pada Klasifikasi Data Penduduk Miskin. *J Rekayasa Teknol Inf.* 2018;2(2):135.
doi:10.30872/jurti.v2i2.1865
11. Nuryaman Y. Komparasi Klasifikasi Penentuan Customer Kredit Menggunakan Algoritma C4.5 Dan Knn Pada Pt Citra Semesta Energy. *J Pilar Nusa Mandiri.* 2018;14(2):233.
doi:10.33480/pilar.v14i2.867
12. Hasan FN, Hikmah N, Utami DY. Perbandingan Algoritma C4.5, KNN, dan Naive Bayes untuk Penentuan Model Klasifikasi Penanggung jawab BSI Entrepreneur Center. *J Pilar Nusa Mandiri.* 2018;14(2):169.
doi:10.33480/pilar.v14i2.908
13. Kurniawan YI. Perbandingan Algoritma Naive Bayes dan C.45 dalam Klasifikasi Data Mining. *J Teknol Inf dan Ilmu Komput.* 2018;5(4):455.
doi:10.25126/jtiik.201854803
14. Rosandy T. Perbandingan Metode Naive Bayes Classifier dengan Metode Decision Tree Untuk Menganalisa Kelancaran Pembiayaan. *J TIM Darmajaya.* 2016;02(01):52-62.
15. Nana D, Elin H. Memilih Metode Penelitian Yang Tepat: Bagi Penelitian Bidang Ilmu Manajemen. *J Ilmu Manaj.* 2018;5(1):288.
<https://jurnal.unigal.ac.id/index.php/ekonologi/article/view/1359>