

## PERBANDINGAN HASIL PREDIKSI KREDIT MACET PADA KOPERASI MENGGUNAKAN ALGORITMA KNN DAN C5.0

Tedi Permana<sup>1\*</sup>), Amril Mutoi Siregar<sup>1)</sup>, Anis Fitri Nur Masruriyah<sup>1)</sup>, Ayu Ratna Juwita<sup>1)</sup>

<sup>1)</sup> Teknik Informatika, Universitas Buana Perjuangan Karawang, Karawang

\*Email Korespondensi : [if16.tedipermana@mhs.ubpkarawang.ac.id](mailto:if16.tedipermana@mhs.ubpkarawang.ac.id)

### ABSTRAK

Koperasi memberikan jasa simpan pinjam dana untuk meningkatkan ekonomi masyarakat, seperti mengembangkan usaha, dan memenuhi kebutuhan hidup sehari-hari. Dalam proses simpan pinjam dana pasti ada permasalahan-permasalahan yang sering terjadi, seperti nasabah terlambat membayar cicilan dana, dan nasabah gagal dalam mengembangkan usahanya, sehingga menyebabkan kredit macet. Pada penelitian ini akan dilakukan prediksi kredit macet untuk meminimalisir terjadinya kredit macet dengan cara memperhatikan data yang dimasukan nasabah untuk proses peminjaman dana. Metode yang digunakan dalam penelitian ini yaitu dengan menerapkan algoritma K-Nearest Neighbor dan algoritma C5.0. Hasil akurasi dari algoritma C5.0 mendapatkan nilai yang lebih baik dari algoritma K-Nearest Neighbor yaitu 86,67%. Sehingga algoritma C5.0 lebih efektif untuk memprediksi kredit macet.

**Kata kunci:** Analisa Kredit, C5.0, K-Nearest Neighbor, Kredit Macet, Koperasi

### ABSTRACT

*Koperasi* provides savings and loan services to improve the community's economy, such as developing businesses and meeting the needs of daily life. In the process of saving and borrowing funds, there must be problems that often occur, such as the customer being late in paying loan installments, and the customer failing to develop his business, causing bad credit. In this research, bad credit predictions will be carried out to minimize the occurrence of bad credit by paying attention to the data entered by the customer for the process of borrowing funds. The method used in this study is to apply the K-Nearest Neighbor algorithm and the C5.0 algorithm. The results of the accuracy of the C5.0 algorithm get a better value than the K-Nearest Neighbor algorithm, namely 86.67%. So that the C5.0 algorithm is more effective at predicting bad credit.

**Keywords:** Bad Credit, C5.0, Credit Analysis, K-Nearest Neighbor, *Koperasi*

### PENDAHULUAN

Koperasi adalah organisasi yang dibuat untuk membantu masyarakat dan para anggotanya dalam urusan keuangan, salah satunya yaitu simpan pinjam dana [1,2]. Masyarakat Indonesia masih tergolong cukup besar dalam hal kredit macet pada peminjaman dana [3]. Permasalahan kredit macet masih menjadi masalah utama di perusahaan pembiayaan, karena akan membuat kondisi keuangan perusahaan tersebut terganggu [4]. Pada umumnya kredit macet di koperasi disebabkan karena pengurus koperasi masih terlalu sederhana dalam melakukan analisa data. Antisipasi yang biasa dilakukan oleh pihak koperasi hanya sebatas pendekatan personal kepada nasabah. Namun antisipasi tersebut masih belum efektif, ketika jumlah nasabah sangat banyak [5]. Salah satu cara untuk menangani masalah tersebut adalah dengan memprediksi kredit macet menggunakan teknik komputasi.

Telah banyak penelitian tentang prediksi, salah satunya penelitian yang dilakukan oleh Vijayarani dan Dhayanand [6], yaitu klasifikasi data untuk prediksi penyakit ginjal. Algoritma yang digunakan yaitu naive bayes dan support vector machine (SVM). Datasets yang digunakan mengandung 584 kasus, dan enam atribut yang digunakan. Hasil yang

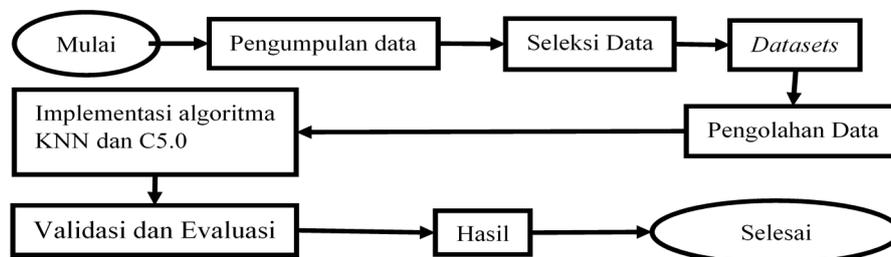
diperoleh menyatakan bahwa SVM lebih akurat dalam klasifikasi data dengan nilai akurasi 76,32% dan dapat membantu tenaga medis dalam proses pemeriksaan penyakit ginjal. Kemudian prediksi kredit macet di bank yang dilakukan oleh Jafar Hamid dan Ahmed [7], algoritma yang digunakan yaitu Bayes Net, J48, dan Naïve Bayes. Datasets yang digunakan yaitu kumpulan data dari perbankan, dengan menggunakan delapan atribut. Hasil yang didapatkan yaitu algoritma J48 lebih baik dalam memprediksi kredit macet dengan nilai akurasi 78,3784%, dan dapat membantu petugas bank dalam menentukan nasabahnya. Selanjutnya penelitian dari Ahmad Fauzi dan Amril Mutoi Siregar [8], yaitu klasifikasi kabupaten kota provinsi Jawa Barat, algoritma yang digunakan yaitu Algoritma Decision Tree. Datasets yang digunakan yaitu database dari pemerintahan bagian statistic dan menggunakan 16 atribut, dengan jumlah kabupaten kota yang ada di Jawa Barat. Hasil yang diperoleh dari penelitian tersebut yaitu mendapatkan nilai akurasi 90%, dan mendapatkan 4 rule yang dijadikan dasar keputusan dalam klasifikasi.

Kemudian penelitian tentang prediksi penyakit jantung yang dilakukan oleh Rohman, Suhartono, dan Supriyanto [9]. Algoritma yang digunakan yaitu, C4.5 dan C4.5 berbasis Adabost. Penelitian ini menggunakan datasets pasien penyakit jantung yang didapatkan dari Universitas California, Irvine (UCI), dan menggunakan 14 atribut. Hasil yang diperoleh dari penelitian tersebut yaitu algoritma C4.5 berbasis adabost mendapatkan nilai akurasi yang lebih tinggi yaitu 92,24%, sehingga praktisi kesehatan dapat menggunakan hasil dari penelitian tersebut sebagai masukan dalam prediksi penyakit jantung. Selanjutnya penelitian oleh Praningsi dan Budi [10] yaitu, prediksi kanker serviks menggunakan algoritma Classification and Regression Tree (CART), Naïve Bayes, dan K-Nearest Neighbor (KNN). Datasets dikumpulkan dari Rumah Sakit Daerah (RSUD) Kediri dan Yayasan Kanker Indonesia Cabang Kediri dan terdapat 12 atribut yang digunakan. Hasil dari penelitian tersebut menyatakan bahwa algoritma Naïve Bayes mampu melakukan klasifikasi dengan baik dengan nilai akurasi 94,44%, sehingga hasil dari penelitian tersebut dapat memberikan keputusan klinis bagi tenaga medis.

Berdasarkan dari teknik – teknik yang telah dilakukan dari penelitian [6,7,8,9,10] terbukti mampu menangani permasalahan yang terjadi. Maka pada penelitian ini akan mengambil judul “Perbandingan Hasil Prediksi Kredit Macet pada Koperasi Menggunakan Algoritma KNN dan C5.0”.

## METODE PENELITIAN

Adapun gambaran umum penelitian yang digunakan, seperti pada Gambar 1. Penelitian dimulai dengan pengumpulan data, seleksi data, pengolahan data dan melakukan klasifikasi (C5.0) juga pengelompokan (KNN). Kemudian proses selanjutnya adalah mengevaluasi algoritma yang digunakan.



Gambar 1. Alur Penelitian

### Pengumpulan Data

Tahap awal yang dilakukan pada penelitian ini adalah dengan mempersiapkan datasets yang diperoleh dari Koperasi Sangkan Mekar. Datasets yang diperoleh berjumlah 120 data

transaksi simpan pinjam dana dari nasabah koperasi. Dari data tersebut terdapat terdapat 106 nasabah dengan keterangan “Lancar”, dan 14 dengan keterangan “Macet” dan terdapat sembilan atribut (RT/RW, Nomor Anggota, Pekerjaan, Status Tinggal, Pendapatan perbulan, Permohonan Pinjaman, Cicilan Bulan Pertama, Jangka Waktu Cicilan, dan Keterangan).

### Seleksi Data

Dari sembilan atribut hanya lima atribut yang digunakan pada proses seleksi data, yaitu status tinggal, penghasilan perbulan, permohonan pinjaman, jangka waktu, dan keterangan. Lima atribut tersebut dipilih karena cocok digunakan untuk perhitungan, dan nilai dari ke lima atribut tersebut lebih bisa digunakan sebagai referensi untuk prediksi kredit macet dibandingkan atribut yang lain. Berikut merupakan beberapa kriteria untuk menentukan seorang nasabah tersebut macet atau lancar dalam proses pembayaran kreditnya:

a. Status Tinggal

Variabel status tinggal dikonversikan dengan sebuah nilai, seperti pada Tabel 1.

Tabel 1. Kriteria Status Tinggal

Nomor	Status Tinggal	Nilai
1	Kost	1
2	KPR	2
3	Milik Orang tua	3
4	MS	4

Pada Tabel 1 dapat dijelaskan untuk status tinggal kost dikonversikan ke dalam nilai 1, untuk KPR dikonversikan ke dalam nilai 2, untuk milik orang tua dikonversikan ke dalam nilai 3, dan untuk MS dikonversikan ke dalam nilai 4.

b. Penghasilan

Variabel penghasilan dikonversikan dalam sebuah nilai ditunjukkan pada Tabel 2.

Tabel 2. Kriteria Penghasilan

Nomor	Pendapatan	Nilai
1	1 juta – 2 juta	1
2	2 juta – 3 juta	2
3	3 juta – 4 juta	3
4	4 juta – 5 juta	4

Pada Tabel 2 dapat dijelas untuk pendapatan 1 juta – 2 juta dikonversikan ke dalam nilai 1, 2 juta – 3 juta dikonversikan ke dalam nilai 2, 3 juta – 4 juta dikonversikan ke dalam nilai 3, dan 4 juta – 5 juta dikonversikan ke dalam nilai 4.

c. Permohonan Pinjaman

Variabel permohonan pinjaman dikonversikan dalam sebuah nilai, seperti Tabel 3.

Tabel 3. Kriteria Permohonan Pinjaman

Nomor	Permohonan Pinjaman	Nilai
1	>5 juta	1
2	4.5 juta	2
3	4 juta	3
4	3 juta	4

Pada Tabel 3 dapat dijelas untuk permohonan pinjaman >5 juta dikonversikan ke dalam nilai 1, 4.5 juta dikonversikan ke dalam nilai 2, 4 juta dikonversikan ke dalam nilai 3, dan 3 juta dikonversikan ke dalam nilai 4.

d. Jangka Waktu Cicilan

Variabel jangka waktu cicilan dikonversikan dengan sebuah nilai, seperti pada Tabel 4

Tabel 4. Kriteria Jangka Waktu Cicilan

Nomor	Jangka Waktu Cicilan	Nilai
1	3 bulan	1
2	6 bulan	2
3	8 bulan	3
4	12 bulan	4

Pada Tabel 4 dapat dijabarkan untuk jangka waktu cicilan 3 bulan dikonversikan ke dalam nilai 1, 6 bulan dikonversikan ke dalam nilai 2, 8 bulan dikonversikan ke dalam nilai 3, dan 12 bulan dikonversikan ke dalam nilai 4.

**Algoritma KNN**

Algoritma KNN merupakan metode yang menggunakan supervised, dimana hasil dari query instance yang baru diklasifikasikan berdasarkan dari kategori pada KNN. Tujuan algoritma KNN yaitu untuk mengklasifikasikan obyek baru berdasarkan atribut dan latih sample. Menurut Suntoro [11] Algoritma KNN mempunyai kelebihan, yaitu memiliki akurasi yang tinggi, insentive terhadap outlier, dan tidak ada asumsi terhadap data. Namun algoritma KNN juga memiliki kekurangan, yaitu perlu untuk memutuskan nilai K optimal, komputasi yang mahal, dan membutuhkan banyak memori. Persamaan untuk mencari Euclidean Distance ditunjukkan pada persamaan (1).

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

Di mana,

- $d(x,y)$  merupakan jarak antara data  $x$  ke data  $y$
- $x_i$  adalah data uji ke  $-i$
- $y_i$  adalah data latih ke  $-i$

Adapun langkah-langkah algoritma KNN sebagai berikut :

1. Tentukan nilai parameter  $k$  (nilai  $k$  dipilih secara manual).
2. Hitung jarak antara data latih dan data uji (metode *Euclidean distance* digunakan sebagian besar peneliti).
3. Urutan data latih berdasarkan jarak terkecil.
4. Menetapkan kelas, dimana kelas yang dipilih adalah kelas dengan jumlah nilai  $k$  terbanyak pada data uji.

**Algoritma C5.0**

Menurut Wijaya, Hasibuan, dan Ramdhani [12] algoritma C5.0 merupakan algoritma klasifikasi dari data mining, dan merupakan salah satu algoritma yang diterapkan pada teknik decision tree . Algoritma C5.0 merupakan algoritma penyempurnan dari ID3 dan C4.5 [13]. Dalam memilih atribut pada pemecahan obyek ada beberapa kelas yang harus dipilih, yaitu atribut yang menghasilkan informasi gain paling besar. Atribut dengan nilai informasi terbesar akan dipilih sebagai parent bagi node selanjutnya [14]. Persamaan untuk membuat pohon keputusan ditunjukkan pada persamaan (2), (3) dan (4).

$$I(S_1, S_2, \dots, S_m) = - \sum_{i=1}^m p_i \log_2 (p_i) \tag{2}$$

Keterangan :

S = himpunan kasus

$S_1$  = jumlah sampel

$P_i$  = proporsi kelas

$$E(A) = \sum_{j=1}^y \frac{S_{1j} + S_{2j}, \dots, S_{mj}}{D} I(S_{1j}, \dots, S_{mj}) \quad (3)$$

$\frac{S_{1j} + S_{2j}, \dots, S_{mj}}{D}$  = jumlah subset j yang dibagi dengan jumlah sampel S.

$$Gain(A) = I(S_1, S_2, \dots, S_m) - E(A) \quad (4)$$

Keterangan :

A = atribut

S = himpunan kasus

$S_1$  = jumlah sampel

## HASIL DAN PEMBAHASAN

### Algoritma KNN

#### a. Menentukan Data Uji dan Data Latih

Pada perhitungan manual KNN data latih yang diambil dari *datassets* yaitu nomor urut 1 samapi 30, karena data dari 31 sampai 120 lebih dominan pada keterangan "Lancar". Maka 30 data sudah cukup untuk mewakili dari 120 data. Sedangkan untuk data uji diambil dari *datassets* nomor urut 31

Tabel 5. Gabungan Data Latih dan Data Uji

No	Status Tinggal	Pendapatan perbulan	Permohonan Pinjaman	Jangka Waktu Cicilan	Keterangan
1	4	3	1	4	Lancar
2	4	3	1	4	Lancar
3	4	3	2	4	Lancar
4	2	3	1	4	Macet
...	...	...	...	...	...
31	4	3	4	2	?

Agar keterangan lancar atau macet dapat diketahui pada setiap nasabah, dari hasil yang sudah diidentifikasi yaitu nilai dari setiap atribut dijumlah. Jika jumlah yang diperoleh < 10 maka keterangannya "Lancar", dan jika > 10 maka keterangannya "Lancar". Dapat dilihat pada Tabel 6 untuk perhitungan Lancar dan Macet.

Tabel 6. Perhitungan Lancar dan Macet

Atribut	Nilai
Status Tinggal	4
Pendapatan perbulan	3
Permohonan Pinjaman	1
Jangka Waktu Cicilan	4
Jumlah	12 (Lancar)

Pada Tabel 6 nilai yang di dapatkan berjumlah 12 maka tergolong ke dalam kategori "lancar", karena nilai 12 lebih dari 10.

#### b. Langkah Penyelesaian

Pada tahap penyelesaian untuk parameter "K" yang digunakan yaitu "4", dan untuk menghitung data baru dengan semua data latih menggunakan *Euclidean Distance*. Pada

perhitungan untuk nomor urut 31 mempunyai nilai yang akan di hitung yaitu 4,3,4,2. Kemudian nilai tersebut dihitung menggunakan rumus *Euclidean Distance* dengan rumus persamaan (1). Contoh perhitungan *Euclidean Distance* bisa dilihat pada Tabel 7.

Tabel 7. Perhitungan *Euclidean Distance*

A	B	C	D	Perhitungan <i>Euclidean Distance</i>
4	3	1	4	$\sqrt{[(4-4)^2+(3-3)^2+(1-4)^2+(4-2)^2]} = 3.6055$
4	3	1	4	$\sqrt{[(4-4)^2+(3-3)^2+(1-4)^2+(4-2)^2]} = 3.6055$
4	3	2	4	$\sqrt{[(4-4)^2+(3-3)^2+(2-4)^2+(4-2)^2]} = 2.8284$
2	3	1	4	$\sqrt{[(2-4)^2+(3-3)^2+(1-4)^2+(4-2)^2]} = 4.1231$
...	...	...	...	...
4	3	4	4	$\sqrt{[(4-4)^2+(3-3)^2+(4-4)^2+(4-2)^2]} \times = 2$

Kemudian untuk tahap selanjutnya yaitu mengurutkan jarak dari data baru dengan data latih dan menentukan tetangga terdekat berdasarkan jarak minimum "K". Selanjutnya menentukan kategori tetangga terdekat dapat dilihat di Tabel 8. Jika kategori  $K < 4$  maka termasuk kategori "Ya", tetapi jika  $K > 4$  maka termasuk kategori "Tidak".

Tabel 8. Kategori Tetangga Terdekat

A	B	C	D	<i>Euclidean Distance</i>	Urutan Jarak	Apakah Termasuk 4-NN	Kategori Ya untuk KNN
4	3	1	4	3.6055	7	Ya ( $K < 4$ )	Lancar
4	3	1	4	3.6055	7	Ya ( $K < 4$ )	Lancar
4	3	2	4	2.8284	5	Ya ( $K < 4$ )	Lancar
2	3	1	4	4.1231	8	Tidak ( $K > 4$ )	-
2	3	1	4	4.1231	8	Tidak ( $K > 4$ )	-
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
4	3	4	4	2	2	Ya ( $K < 4$ )	Lancar

Hasil Akurasi yang diperoleh dari dari Tabel 8 kategori tetangga terdekat yaitu 90%:

$$\frac{27}{30} \times 100\% = 90\%$$

Berdasarkan Tabel 8 terdapat 27 data yang sesuai dengan *datasets*, yaitu 26 untuk "Lancar" dan 1 untuk "Macet". Jadi dapat disimpulkan untuk data nomor urut 31 termasuk ke dalam kategori "Lancar",

Tabel 9. Hasil Prediksi K-NN untuk *Datasets* nomor urut 31

No	Status Tinggal	Pendapatan perbulan	Permohonan Pinjaman	Jangka Waktu Cicilan	Keterangan
31	4	3	4	2	Lancar

Pada pengujian menggunakan aplikasi Rapidminer algoritma KNN mendapatkan hasil akurasi 83.33%, seperti pada Tabel 10.

Tabel 10. Hasil KNN Menggunakan *Rapid Miner*

	True Lancar	True Macet	Class Percision
Prediksi Lancar	28	4	86,67%
Prediksi Macet	1	0	0,00%
<i>Class recall</i>	96,30%	0,00%	
Akurasi	83,33% ± 17,57% ( <i>micro average</i> :83,87%)		

### Algoritma C5.0

Pada proses perhitungan manual ini menggunakan 30 data dari 120 *datasets*, dengan jumlah anggota 26 orang "Lancar", dan 4 orang "Macet".

#### 1. Menghitung jumlah *entropy Node 1*

$$\text{Entropy total} \left( -\frac{26}{30} \times \log_2 \left( \frac{26}{30} \right) + \left( -\frac{4}{30} \times \log_2 \left( \frac{4}{30} \right) \right) \right) = 0,56649$$

Entropy (ST)

- $\text{entropy (MS)} = \left( -\frac{23}{24} \times \log_2 \left( \frac{23}{24} \right) + \left( -\frac{1}{24} \times \log_2 \left( \frac{1}{24} \right) \right) \right) = 0,24988$
- $\text{entropy (KPR)} = \left( -\frac{2}{4} \times \log_2 \left( \frac{2}{4} \right) + \left( -\frac{2}{4} \times \log_2 \left( \frac{2}{4} \right) \right) \right) = 1$
- $\text{entropy (MOT)} = \left( -\frac{1}{1} \times \log_2 \left( \frac{1}{1} \right) + \left( -\frac{0}{1} \times \log_2 \left( \frac{0}{1} \right) \right) \right) = 0$
- $\text{entropy (KOS)} = \left( -\frac{0}{1} \times \log_2 \left( \frac{0}{1} \right) + \left( -\frac{1}{1} \times \log_2 \left( \frac{1}{1} \right) \right) \right) = 0$

Entropy (PP)

- $\text{entropy (1-2 juta)} = \left( -\frac{2}{2} \times \log_2 \left( \frac{2}{2} \right) + \left( -\frac{0}{2} \times \log_2 \left( \frac{0}{2} \right) \right) \right) = 0$
- $\text{entropy (2-3 juta)} = \left( -\frac{3}{3} \times \log_2 \left( \frac{3}{3} \right) + \left( -\frac{0}{3} \times \log_2 \left( \frac{0}{3} \right) \right) \right) = 0$
- $\text{entropy (3-4 juta)} = \left( -\frac{21}{25} \times \log_2 \left( \frac{21}{25} \right) + \left( -\frac{4}{25} \times \log_2 \left( \frac{4}{25} \right) \right) \right) = 0,6342$

Entropy (PPin)

- $\text{entropy (3 juta)} = \left( -\frac{6}{6} \times \log_2 \left( \frac{6}{6} \right) + \left( -\frac{0}{6} \times \log_2 \left( \frac{0}{6} \right) \right) \right) = 0$
- $\text{entropy (4 juta)} = \left( -\frac{8}{8} \times \log_2 \left( \frac{8}{8} \right) + \left( -\frac{0}{8} \times \log_2 \left( \frac{0}{8} \right) \right) \right) = 0$
- $\text{entropy (4,5 juta)} = \left( -\frac{5}{6} \times \log_2 \left( \frac{5}{6} \right) + \left( -\frac{1}{6} \times \log_2 \left( \frac{1}{6} \right) \right) \right) = 0,64999$
- $\text{entropy (5 juta)} = \left( -\frac{7}{10} \times \log_2 \left( \frac{7}{10} \right) + \left( -\frac{3}{10} \times \log_2 \left( \frac{3}{10} \right) \right) \right) = 0,88132$

Entropy (JW)

- $\text{entropy (12 Bulan)} = \left( -\frac{17}{20} \times \log_2 \left( \frac{17}{20} \right) + \left( -\frac{3}{20} \times \log_2 \left( \frac{3}{20} \right) \right) \right) = 0,60987$
- $\text{entropy (8 Bulan)} = \left( -\frac{7}{7} \times \log_2 \left( \frac{7}{7} \right) + \left( -\frac{0}{7} \times \log_2 \left( \frac{0}{7} \right) \right) \right) = 0$
- $\text{entropy (6 Bulan)} = \left( -\frac{1}{2} \times \log_2 \left( \frac{1}{2} \right) + \left( -\frac{1}{2} \times \log_2 \left( \frac{1}{2} \right) \right) \right) = 1$
- $\text{entropy (3 Bulan)} = \left( -\frac{1}{1} \times \log_2 \left( \frac{1}{1} \right) + \left( -\frac{0}{1} \times \log_2 \left( \frac{0}{1} \right) \right) \right) = 0$

#### 2. Menghitung Jumlah Gain

- $\text{Gain (ST)} = 1 - \left( \frac{4}{4} \times 1 \right) = 0$
- $\text{Gain (PP)} = 1 - \left( \frac{4}{4} \times 1 \right) = 0$
- $\text{Gain (PPin)} = 1 - \left( \frac{3}{4} \times 0,91833 \right) + \left( \frac{1}{4} \times 0 \right) = 0,31125$
- $\text{Gain (JW)} = 1 - \left( \frac{3}{4} \times 0,91833 \right) + \left( \frac{1}{4} \times 0 \right) = 0,31125$

Dari perhitungan entropy diatas maka diperoleh hasil seperti pada Tabel 11.

Tabel 11. *Node 1.1* (KPR)

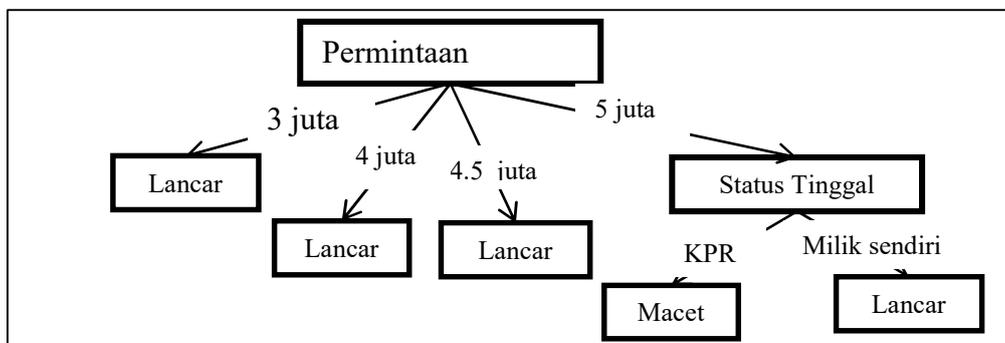
<i>Node</i>			Jumlah	Lancar	Macet	Entropy	Gain
1.1	Total		4	2	2	1	
	ST						0
		KPR	4	2	2	1	
	PP						0
		3-4 juta	4	2	2	1	
	PPin						0,31125
		4,5 juta	1	1	0	0	
		5 juta	3	1	2	0,91833	
	JW						0,31125
		12 Bulan	3	1	2	0,91833	
		6 Bulan	2	2	0	0	

Keterangan :

- ST = Status Tinggal
- PP = Penghasilan perbulan
- PPin = Permintaan Pinjaman
- JW = Jangka Waktu
- MS = Milik Sendiri
- MOT = Milik Orang Tua
- KOS = Kost
- KPR = Kredit Pemilikan Rumah

Dari Tabel 11 maka dapat dilihat yang mempunyai nilai gain tertinggi yaitu PPin dan JW. Maka PPin dan JW dapat dijadikan sebagai akar dari pohon keputusan *Node 1.1* .Dari pohon keputusan *Node 1.1* masih ada yang belum mempunyai hasil akhir, yaitu PPin => 5 juta, dan JW => 12 bulan. Maka harus dihitung kembali untuk mencari *Node 1.2* dan *Node 1.3*. lakukan perhitungan dengan cara yang sama sampai semua atribut mempunyai hasil akhir.

### Hasil Akhir Pohon Keputusan



Gambar 2. Hasil Akhir Pohon Keputusan.

Dari Gambar 2 dapat dijelaskan, jika permintaan pinjaman (PPin) 3 juta = "Lancar", 4 juta = "Lancar", 4.5 juta = "Lancar". Sedangkan untuk PPin 5 juta harus masuk dulu ke atribut Status Tinggal (ST), jika status tinggal KPR = "Macet", sedangkan milik sendiri = "Lancar". Pada pengujian menggunakan tools Rapidminer algoritma C5.0 mendapatkan hasil akurasi 86.67%, seperti pada Tabel 12.

Tabel 12. Hasil Akurasi C5.0 Rapidminer

	<b>True Lancar</b>	<b>True Macet</b>	<b>Class Percision</b>
<b>Prediksi Lancar</b>	26	4	86,67%
<b>Prediksi Macet</b>	0	0	0,00%
<b>Class recall</b>	100%	0,00%	
<b>Akurasi</b>	86,67% ± 17,21% (micro average : 86,67%)		

## KESIMPULAN

Berdasarkan hasil perhitungan dengan algoritma KNN dan C5.0 dengan datasets yang digunakan sebanyak 30 data transaksi simpan pinjam, dan lima atribut yang digunakan, nilai akurasi yang diperoleh algoritma C5.0 lebih baik dari KNN yaitu 86.67%. Sedangkan untuk algoritma KNN mendapatkan nilai akurasi 83.33%, sehingga algoritma C5.0 lebih cocok digunakan sebagai patokan untuk prediksi kredit macet di koperasi.

## UCAPAN TERIMA KASIH

Naskah ilmiah ini adalah sebagian dari penelitian Tugas Akhir milik Tedi Permana yang dibimbing oleh Amril Mutoi Siregar dan Anis Fitri Nur Masruriyah.

## REFERENSI

- [1] KBBI. (2019). Perpustakaan.
- [2] Juniantara, I. W., & Riana, I. G. (2015). Pengaruh Motivasi dan Kepuasan Kerja terhadap Kinerja Karyawan Koperasi di Denpasar. E-Jurnal Ekonomi dan Bisnis, 09(4), 611-628.
- [3] Sucipto, A. (2015). Pada Koperasi Simpan Pinjam Dengan Menggunakan. Jurnal DISPROTEK, 6(1), 75-87. <https://doi.org/10.34001/jdpt.v6i1.200>
- [4] Fauzi, A. (2018). Kredit macet , npl dan pengaruhnya terhadap kinerja perusahaan pada perusahaan pembiayaan 1. 2, 27-36.
- [5] Sucipto, A. (2015). Pada Koperasi Simpan Pinjam Dengan Menggunakan. Jurnal DISPROTEK, 6(1), 75-87. <https://doi.org/10.34001/jdpt.v6i1.200>
- [6] Vijayarani, S., & Dhayanand, S. (2015). Data Mining Classification Algorithms for Kidney Disease Prediction. International Journal on Cybernetics & Informatics, 4(4), 13-25. <https://doi.org/10.5121/ijci.2015.4402>
- [7] Jafar Hamid, A., & Ahmed, T. M. (2016). Developing Prediction Model of Loan Risk in Banks Using Data Mining. Machine Learning and Applications: An International Journal, 3(1), 1-9. <https://doi.org/10.5121/mlaj.2016.3101>
- [8] Fauzi, A., & M.S, A. (2020). Klasifikasi Kabupaten Kota Provinsi Jawa Barat Berdasarkan Pendapatan Dari Sektor Pertanian Dengan Algoritma Decision Tree. 13(1), 1-8. <https://doi.org/10.30998/faktorexacta.v13i1.5542>
- [9] Rohman, A., Suhartono, V., & Supriyanto, C. (2017). Penerapan algoritma c4.5 berbasis. 13, 13-19.
- [10] Praningki, T., & Budi, I. (2017). Sistem Prediksi Penyakit Kanker Serviks Menggunakan CART , Naive Bayes , dan k-NN. 4(2).

- [11] Suntoro, J. (2019). *Data Mining : Algoritma dan Implementasi dengan Pemrograman PHP*. Jakarta: PT Elex Media Komputindo.
- [12] udarma, I. M. (2018). Implementasi Algoritma C5 . 0 pada Penilaian. *Majalah Ilmiah Teknologi Elektro*, 17(3), 1-6.
- [13] Hanani, A., Dachlan, H., & Santoso, P. (2015). Pemberian Harakat Bahasa Arab Menggunakan Metode N-Gram Dan C5.0. *Jurnal EECCIS*, 9(1), pp.73-78.
- [14] Enterprise, J. (2018). *Pemrograman Database dengan Python dan Mysql*. Elex Media Komputindo.