

# Klasifikasi Permasalahan Kredit Macet Pada Bank Menggunakan Algoritma *Decision Tree* C4.5

Anida Nirwana  
Universitas Buana Perjuangan  
Karawang, Indonesia  
If17.anidanirwana@mhs.ubpkarawang.ac.id

Amril Mutoi Siregar  
Universitas Buana Perjuangan  
Karawang, Indonesia  
amrilmutoi@ubpkarawang.ac.id

Rahmat  
Universitas Buana Perjuangan  
Karawang, Indonesia  
Rahmat@ubpkarawang.ac.id

## Abstract—

Dalam pemberian kredit tentunya lembaga keuangan harus merencanakan sedemikian rupa dan berusaha mengurangi risiko permasalahan adanya kredit macet. Faktor yang sering terjadi dalam permasalahan kredit macet ini utamanya dari pihak nasabah, karena kegagalan bisnis ditambah lagi dengan adanya pandemi *covid-19* dan juga faktor yang sering terjadi karena ketidaktelitiannya pihak bank dalam menganalisis data calon nasabah yang memiliki karakter yang tidak baik serta saat melakukan analisis kelayakan usaha nasabah pengetahuan pihak bank terbatas, sehingga analisis kredit tidak tepat. Data penelitian ini diambil dari web kaggle pada bulan November tahun 2020 sebanyak 10.127 data dan 10 variabel. Penelitian ini bertujuan untuk mengurangi risiko permasalahan kredit macet dengan cara mengklasifikasikan permasalahan kredit macet pada bank dengan menerapkan algoritma decision tree c4.5. Dalam mengklasifikasikan kredit macet agar menghasilkan pohon keputusan yang akan menjadi penunjang dalam proses perhitungan tingkat keakurasian data. Dataset dibagi menjadi dua yakni data *training* dan data *testing* dengan pembagian data 60 : 40. Pengujian ini menggunakan excel untuk perhitungan manual dengan model data *training*, python dengan model data *testing* dan menggunakan tool weka 3.8.5 dengan model data *testing* dengan pembagian data 60 : 40, untuk 60% yaitu data *training* dan 40% data *testing*. Didalam pembagian data untuk data *training* yaitu berjumlah 6076 data dan untuk data *testing* berjumlah 4051 data dengan memiliki nilai *accuracy* sebesar 99,9753%, *precision* sebesar 100%, *recall* sebesar 99,8%, dan *f-measure* sebesar 99,9%.

**Kata kunci** — Kredit Macet, Algoritma C4.5, python

## I. PENDAHULUAN

Kredit memiliki peran yang sangat penting dalam memenuhi kebutuhan perekonomian masyarakat, karena lembaga keuangan memiliki kebijakan dari pemerintah dalam perkembangan perekonomian di Indonesia. Dalam pemberian kredit tentunya lembaga keuangan harus merencanakan sedemikian rupa dan berusaha mengurangi risiko permasalahan dalam pemberian kredit. Faktor dari permasalahan kredit ini dapat mengakibatkan berkurangnya keuntungan dan modal bagi bank tersebut [1]. Faktor yang sering terjadi dalam permasalahan kredit macet ini utamanya dari pihak nasabah, karena kegagalan bisnis ditambah lagi dengan adanya pandemi *covid-19* dan juga faktor yang sering terjadi karena ketidaktelitiannya pihak bank dalam menganalisis data calon nasabah yang memiliki karakter yang tidak baik serta saat melakukan analisis kelayakan usaha nasabah pengetahuan pihak bank terbatas, sehingga analisis kredit tidak tepat. Hal ini yang mengakibatkan tingginya permasalahan kredit macet. Melalui penelitian ini diharapkan dapat membantu pihak lembaga keuangan dalam membaca pola pembayaran dari nasabah seperti kemampuan nasabah dalam bidang bisnis, jaminan yang akan diberikan nasabah kepada pihak bank, dan sifat atau watak nasabah diluar ketentuan pihak bank yaitu BI *Checking*. BI *Checking* merupakan istilah *output* yang dihasilkan oleh sistem informasi debitur (SID), didalam BI *checking* terdapat 5 kategori yang akan menjadi ketentuan pihak bank dalam pemberian kredit kepada nasabah diantaranya kredit lancar, kredit dalam perhatian khusus (DPK), kredit tidak lancar, kredit diragukan, dan kredit macet. Didalam penelitian ini teknik yang digunakan yaitu teknik data mining klasifikasi dengan metode *decision tree* C4.5 untuk menentukan apakah nasabah tersebut layak mendapatkan kredit atau tidak.

Penelitian tentang klasifikasi telah banyak ditemukan, dalam penelitian yang membahas tentang perbandingan Algoritma *Decision Tree* adalah penelitian yang dilakukan oleh Kamil, Widiyanto dan Chamidah [2] tentang Perbandingan Metode Decision Tree Dengan Naïve Bayes Dalam Klasifikasi Tumor Otak Citra MRI. *Decision Tree* mempunyai nilai yang lebih tinggi dengan *accuracy* 96%, *spesificaty* 96%, dan *sensitivity* 96%, sedangkan metode *Naïve Bayes* memiliki nilai *accuracy* 91%, *spesificaty* 90%, dan *sensitivity* 93%. Selain itu penelitian yang dilakukan oleh Wahyuningsih dan Utari [3] tentang Perbandingan Metode K-Nearest Neighbor, Naïve Bayes dan Decision Tree untuk Prediksi Kelayakan Pemberian Kredit. Hasil dari penelitian ini algoritma *Decision Tree* memperoleh nilai akurasi sebesar 92,21%, algoritma *Naïve Bayes* memiliki akurasi sebesar 81,83% dan algoritma K-NN memiliki akurasi sebesar 81,82%. Maka dapat disimpulkan bahwa algoritma *Decision Tree* layak untuk digunakan karena memiliki tingkat akurasi tertinggi. Penelitian yang dilakukan oleh Hana [4] tentang Klasifikasi Penderita Penyakit Diabetes Menggunakan Algoritma Decision Tree C4.5. Hasil yang didapatkan melalui pengujian dengan cara membentuk pohon keputusan memiliki akurasi sebesar 97,12%, *precision* sebesar 93,02%, dan *recall* sebesar 100,00%. Penelitian lainnya oleh Karyadiputra dan Hijriana [5] tentang Penerapan Algoritma Decision Tree C4.5 Untuk Klasifikasi Penentuan Daftar Prioritas Pengembangan Jembatan. Hasil dari penelitian dengan menggunakan algoritma *Decision Tree* menghasilkan *performance* akurasi sebesar 82.84% dengan AUC sebesar 0,825 sehingga termasuk dalam kategori klasifikasi terbaik.

Berdasarkan hasil yang diperoleh dari penelitian diatas dengan menggunakan algoritma *decision tree* c4.5 mendapatkan nilai akurasi yang tinggi, maka dari itu peneliti memilih algoritma *decision tree* c4.5 untuk mengklasifikasikan dan mengetahui hasil akurasi dari penelitian dengan judul “Klasifikasi Permasalahan Kredit Macet Pada Bank Menggunakan Algoritma *Decision Tree* C4.5”.

## II. DATA DAN METODE

### A. Kredit Macet

Kredit macet merupakan pembayaran pinjaman yang mengalami kesulitan dalam mengembalikan tagihan kepada bank sesuai dengan ketentuan yang sudah disepakati [6]. Kredit macet berdampak merugikan bagi bank, dimana kerugian yang disebabkan karena tidak menerima dana yang disalurkan dan bunga yang tidak dapat diterima mengakibatkan pendapatan bunga dan total pendapatan bank akan berkurang [7].

### B. Algoritma Decision Tree C4.5

Algoritma *decision tree* merupakan salah satu metode klasifikasi dan prediksi yang terbaik untuk digunakan dan mudah dipahami. Salah satu pengembangan dari algoritma *decision tree* adalah C4.5. Algoritma C4.5 sering digunakan dalam penelitian yang dilakukan karena menghasilkan beberapa aturan dan pohon keputusan yang bertujuan untuk meningkatkan akurasi prediksi yang dilakukan [8]. Menurut Darmawan [9] untuk membangun pohon keputusan, algoritma *decision tree* ini memiliki rincian secara umum, antara lain :

1. Pilih atribut sebagai *root*-nya
2. Buat cabang untuk setiap nilai.
3. Bagi kasus dalam simpulnya.
4. Ulangi proses pada cabang sampai kasus mendapatkan kelas yang sama.

Dalam perhitungan algoritma *decision tree* C4.5 terdapat dua langkah, diantaranya :

1. Menghitung nilai *entropy*

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

Keterangan :

S : Himpunan Kasus

N : Jumlah Partisi S

Pi : Proporsi dari Si terhadap S

2. Menghitung *information Gain*

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

Keterangan :

S : Himpunan Kasus

A : Atribut

N : Jumlah partisi atribut A

|Si| : Jumlah kasus pada partisi ke-i

|S| : Jumlah kasus pada S

### C. Confusion Matrix

*Confusion matrix* merupakan suatu metode yang biasa digunakan untuk menghitung akurasi di bidang data mining. Menurut Paskalis [10] *Confusion matrix* ini akan melakukan kalkulasi yaitu melakukan empat keluaran :

1. *Recall*, jumlah data positif yang teridentifikasi benar. *Recall* dihitung dengan cara membagi data benar yang bernilai *True Positive* (TP) dengan hasil penjumlahan data benar yang bernilai *True Positive* (TP) dan data salah yang bernilai *False Negative* (FN).

$$recall = \frac{TP}{TP + FN} \quad (3)$$

2. *Precision*, proporsi kasus dengan hasil positif yang benar. nilai *precision* dihitung dengan cara membagi jumlah data benar yang bernilai *True Positive* (TP) dengan hasil penjumlahan data benar yang bernilai *True Positive* (TP) dan data salah yang bernilai *False Positive* (FP).

3. 
$$Precision = \frac{TP}{TP + FP} \quad (4)$$

4. *Accuracy*, perbandingan kasus yang teridentifikasi benar dengan jumlah seluruh kasus. Nilai *accuracy* dihitung dengan cara membagi jumlah data yang diklasifikasi secara benar dengan total sampel data *testing* yang diuji.

$$Accuracy = \frac{\text{jumlah prediksi benar}}{\text{jumlah total prediksi}} \quad (5)$$

5. *F-Measure*, nilai *F-Measure* didapat dari perhitungan pembagian hasil dari perkalian *precision* dan *recall* dengan hasil penjumlahan *precision* dan *recall*, kemudian dikalikan dua.

(6)

$$F - Measure = 2 * \frac{precision * recall}{precision + recall}$$

Tabel 1 Confusion Matrix [9]

Classification	Predicted Class	
	Class = Yes	Class = No
Class = Yes	a (True Positive – TP)	b (False Negative – FN)
Class = No	c (False Positive – FP)	d (True Negative – TN)

D. Python

Python merupakan bahasa pemrograman tingkat tinggi yang dikembangkan pada tahun 1989 oleh Guido van Rossum dan pertama kali dirilis pada tahun 1991 [11]. Didalam bahasa pemrograman python terdapat beberapa fitur antara lain[11] :

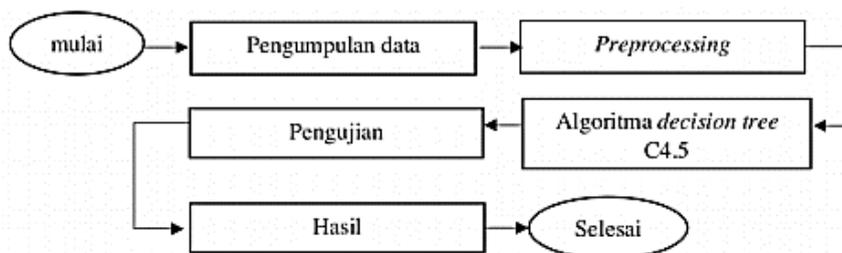
1. Memiliki *library* yang luas
2. Memiliki *tata bahasa* yang jelas dan mudah dipahami.
3. Memiliki aturan *layout source code* yang memudahkan dalam melakukan pemeriksaan, membaca ulang dan menulis ulang *source code* tersebut.
4. Memiliki modular yang dapat dikembangkan dengan membuat modul baru baik dibuat dalam bahasa python atau C/C++.
5. Memiliki fungsi pengumpulan sampah secara otomatis

E. Weka

Weka merupakan *software* yang digunakan untuk mengaplikasikan beberapa proses terkait data *mining* dengan berbagai algoritma *machine learning* mulai dari *classification*, *clustering*, *association rule*, *visualization*, dan *pre-processing* yang dikembangkan di Universitas Waikato, Selandia Baru [12].

F. Gambaran Umum Penelitian

Adapun gambaran umum penelitian yang digunakan.



Gambar 1 Gambaran Umum Penelitian

G. Pengumpulan Data

Penelitian ini dimulai dengan melakukan pengumpulan data melalui salah satu *online repository* dan mendapatkan *dataset* tentang permasalahan kredit macet yang diakses pada situs [www.kaagle.com/sakshigoyal7/credit-card-customer](http://www.kaagle.com/sakshigoyal7/credit-card-customer) pada bulan November tahun 2020 yang terdiri dari 10.127 data dan 10 atribut.

H. Preprocessing

Didalam tahapan *preprocessing* data yaitu melakukan seleksi atribut yang akan digunakan untuk memilih atribut terbaik.

III. HASIL DAN PEMBAHASAN

A. Hasil Penelitian

*Dataset* permasalahan kredit macet ini merupakan dataset yang sebelum dilakukan *preprocessing* diambil dari salah satu situs web lebih tepatnya *website* [kaggle.com](http://kaggle.com) pada bulan November tahun 2020 dengan jumlah data sebanyak 10.127 data dan 10 variabel

Setelah data terkumpul tahapan selanjutnya yaitu tahap *preprocessing* melakukan pembersihan dan seleksi atribut untuk mendapatkan atribut-atribut yang bernilai relevan diantaranya usia, jumlah tanggungan, kategori pendapatan, total bulan dan satu atribut *predictor* yaitu keterangan. *Dataset* dibagi menjadi dua yakni data *training* dan data *testing* dengan pembagian

data 60 : 40, dimana 60% data *training* dan 40% data *testing*. Didalam pembagian jumlah datanya untuk data *training* berjumlah 6.076 data dan untuk data *testing* berjumlah 4.051 data. Pada algoritma *decision tree* c4.5 pengujian dilakukan dengan tiga tahapan menggunakan excel, weka, dan python.

Tabel 2 Data *Training* Dan Data *Testing*

No	Usia	Jumlah Tanggungan	Kategori Pendapatan	Total Bulan	Usia	Jumlah Tanggungan	Kategori Pendapatan	Total Bulan	Keterangan
1.	51	2	Rp5.765.180	43	Diragukan	Lancar	Diragukan	Lancar	Tidak Layak
2.	55	3	Rp5.765.180	48	Diragukan	Lancar	Diragukan	Lancar	Tidak Layak
3.	42	5	Rp7.205.620	37	Lancar	Diragukan	Diragukan	Lancar	Tidak Layak
4.	53	3	Rp8.646.750	33	Diragukan	Lancar	Lancar	Diragukan	Tidak Layak
5.	43	3	Rp13.113.225	33	Lancar	Lancar	Lancar	Diragukan	Tidak Layak
6.	51	1	Rp8.646.750	43	Diragukan	Lancar	Lancar	Lancar	Tidak Layak
7.	56	3	Rp13.113.225	44	Diragukan	Lancar	Lancar	Lancar	Tidak Layak
...	...	...	...	...	...	...	...	...	...
...	.....	...	...	...	...	...	...	...	...
6076	41	4	Rp7.205.620	36	Lancar	Diragukan	Diragukan	Lancar	Tidak Layak

1. Hasil Pengujian Algoritma Decision Tree Menggunakan Excel

a. Menghitung entropy node 1

Node 1	Jumlah Kasus (S)	Layak (S1)	Tidak Layak (S2)	Entropy	Gain	
<b>Total</b>	<b>6076</b>	<b>793</b>	<b>5283</b>	<b>0,558844769</b>		
Usia	Lancar	4346	793	3553	0,685443495	0,068565403
	Diragukan	1730	0	1730	0	
Jumlah Tanggungan	Lancar	4419	793	3626	0,678877557	0,065105479
	Diragukan	1657	0	1657	0	
Kategori Pendapatan	Lancar	2580	793	1787	0,89010758	0,180885988
	Diragukan	3496	0	5351	0	
Total Bulan	Lancar	4300	793	3507	0,689648489	0,070778854
	Diragukan	1776	0	1776	0	

Gambar 2 node 1

$$\text{Entropy (total)} = \left(-\frac{793}{6076}\right) * \text{Imlog}_2\left(\frac{793}{6076}\right) + \left(-\frac{5283}{6076}\right) * \text{Imlog}_2\left(\frac{5283}{6076}\right) = 0,558844769$$

**Entropy (usia)**

$$\text{Lancar} = \left(-\frac{793}{4346}\right) * \text{Imlog}_2\left(\frac{793}{4346}\right) + \left(-\frac{3553}{4346}\right) * \text{Imlog}_2\left(\frac{3553}{4346}\right) = 0,685443495$$

$$\text{Diragukan} = \left(-\frac{0}{1730}\right) * \text{Imlog}_2\left(\frac{0}{1730}\right) + \left(-\frac{1730}{1730}\right) * \text{Imlog}_2\left(\frac{1730}{1730}\right) = 0$$

**Entropy (Jumlah Tanggungan)**

$$\text{Lancar} = \left(-\frac{793}{4419}\right) * \text{Imlog}_2\left(\frac{793}{4419}\right) + \left(-\frac{3626}{4419}\right) * \text{Imlog}_2\left(\frac{3626}{4419}\right) = 0,678877557$$

$$\text{Diragukan} = \left(-\frac{0}{1657}\right) * \text{Imlog}_2\left(\frac{0}{1657}\right) + \left(-\frac{1657}{1657}\right) * \text{Imlog}_2\left(\frac{1657}{1657}\right) = 0$$

**Entropy (Kategori Pendapatan)**

$$\text{Lancar} = \left(-\frac{793}{2580}\right) * \text{Imlog}_2\left(\frac{793}{2580}\right) + \left(-\frac{1787}{2580}\right) * \text{Imlog}_2\left(\frac{1787}{2580}\right) = 0,89010758$$

$$\text{Diragukan} = \left(-\frac{0}{3496}\right) * \text{Imlog}_2\left(\frac{0}{3496}\right) + \left(-\frac{3496}{3496}\right) * \text{Imlog}_2\left(\frac{3496}{3496}\right) = 0$$

**Entropy (Total Bulan)**

$$\text{Lancar} = \left(-\frac{793}{4300}\right) * \text{Imlog}_2\left(\frac{793}{4300}\right) + \left(-\frac{3507}{4300}\right) * \text{Imlog}_2\left(\frac{3507}{4300}\right) = 0,689648489$$

$$\text{Diragukan} = \left(-\frac{0}{1776}\right) * \text{Imlog}2\left(\frac{0}{1776}\right) + \left(-\frac{1776}{1776}\right) * \text{Imlog}2\left(\frac{1776}{1776}\right) = 0$$

b. Menghitung Gain Node 1

$$\text{Gain (Total, Usia)} = 0,558844769 - \left(\frac{4346}{6076} * 0,685443495\right) + \left(\frac{1730}{6076} * 0\right) = 0,068565403$$

$$\text{Gain (Total, Jumlah Tanggungan)} = 0,558844769 - \left(\frac{4419}{6076} * 0,678877557\right) + \left(\frac{1657}{6076} * 0\right) = 0,065105479$$

$$\text{Gain (Total, Kategori Pendapatan)} = 0,558844769 - \left(\frac{2580}{6076} * 0,89010758\right) + \left(\frac{3496}{6076} * 0\right) = 0,180885988$$

$$\text{Gain (Total, Total Bulan)} = 0,558844769 - \left(\frac{4300}{6076} * 0,689648489\right) + \left(\frac{1776}{6076} * 0\right) = 0,070778854$$

c. Menghitung Entropy Node 1.1

Node 1.1		Jumlah Kasus (S)	Layak (S1)	Tidak Layak (S2)	Entropy	Gain
Kategori Pendapatan Lancar		2580	793	1787	0,89010758	
Usia	Lancar	1837	793	1044	0,986490693	0,187710912
	Diragukan	743	0	743	0	
Jumlah Tanggungan	Lancar	2001	793	1208	0,968746166	0,138766076
	Diragukan	579	0	579	0	
Total Bulan	Lancar	1860	793	1067	0,984289086	0,18050382
	Diragukan	720	0	720	0	

Gambar 3 Node 1.1

$$\text{Entropy (Kategori pendapatan (Lancar))} = \left(-\frac{793}{2580}\right) * \text{Imlog}2\left(\frac{793}{2580}\right) + \left(-\frac{1787}{2580}\right) * \text{Imlog}2\left(\frac{1787}{2580}\right) = 0,89010758$$

**Entropy (usia)**

$$\text{Lancar} = \left(-\frac{793}{1837}\right) * \text{Imlog}2\left(\frac{793}{1837}\right) + \left(-\frac{1044}{1837}\right) * \text{Imlog}2\left(\frac{1044}{1837}\right) = 0,986490693$$

$$\text{Diragukan} = \left(-\frac{0}{743}\right) * \text{Imlog}2\left(\frac{0}{743}\right) + \left(-\frac{743}{743}\right) * \text{Imlog}2\left(\frac{743}{743}\right) = 0$$

**Entropy (Jumlah Tanggungan)**

$$\text{Lancar} = \left(-\frac{793}{2001}\right) * \text{Imlog}2\left(\frac{793}{2001}\right) + \left(-\frac{1208}{2001}\right) * \text{Imlog}2\left(\frac{1208}{2001}\right) = 0,968746166$$

$$\text{Diragukan} = \left(-\frac{0}{579}\right) * \text{Imlog}2\left(\frac{0}{579}\right) + \left(-\frac{579}{579}\right) * \text{Imlog}2\left(\frac{579}{579}\right) = 0$$

**Entropy (Total Bulan)**

$$\text{Lancar} = \left(-\frac{793}{1860}\right) * \text{Imlog}2\left(\frac{793}{1860}\right) + \left(-\frac{1067}{1860}\right) * \text{Imlog}2\left(\frac{1067}{1860}\right) = 0,984289086$$

$$\text{Diragukan} = \left(-\frac{0}{720}\right) * \text{Imlog}2\left(\frac{0}{720}\right) + \left(-\frac{720}{720}\right) * \text{Imlog}2\left(\frac{720}{720}\right) = 0$$

d. Menghitung Gain 1.1

$$\text{Gain (Total, Usia)} = 0,89010758 - \left(\frac{1837}{2580} * 0,986490693\right) + \left(\frac{743}{2580} * 0\right) = 0,187710912$$

$$\text{Gain (Total, Jumlah Tanggungan)} = 0,89010758 - \left(\frac{2001}{2580} * 0,968746166\right) + \left(\frac{579}{2580} * 0\right) = 0,138766076$$

$$\text{Gain (Total, Total Bulan)} = 0,89010758 - \left(\frac{1860}{2580} * 0,984289086\right) + \left(\frac{720}{2580} * 0\right) = 0,18050382$$

e. Menghitung entropy 1.2

Node 1.2		Jumlah Kasus (S)	Layak (S1)	Tidak Layak (S2)	Entropy	Gain
Usia (Lancar)		1837	793	1044	0,986490693	
Jumlah Tanggungan	Lancar	1328	793	535	0,972599843	0,283380954
	Diragukan	509	0	509	0	
Total Bulan	Lancar	1136	793	343	0,883645341	0,440044799
	Diragukan	701	0	701	0	

Gambar 4 Node 1.2

$$\text{Entropy (Usia (Lancar))} = \left(-\frac{793}{1837}\right) * \text{Imlog}_2\left(\frac{793}{1837}\right) + \left(-\frac{1044}{1837}\right) * \text{Imlog}_2\left(\frac{1044}{1837}\right) = 0,986490693$$

**Entropy (Jumlah Tanggungan)**

$$\text{Lancar} = \left(-\frac{793}{1328}\right) * \text{Imlog}_2\left(\frac{793}{1328}\right) + \left(-\frac{535}{1328}\right) * \text{Imlog}_2\left(\frac{535}{1328}\right) = 0,972599843$$

$$\text{Diragukan} = \left(-\frac{0}{509}\right) * \text{Imlog}_2\left(\frac{0}{509}\right) + \left(-\frac{509}{509}\right) * \text{Imlog}_2\left(\frac{509}{509}\right) = 0$$

**Entropy (Total Bulan)**

$$\text{Lancar} = \left(-\frac{793}{1136}\right) * \text{Imlog}_2\left(\frac{793}{1136}\right) + \left(-\frac{343}{1136}\right) * \text{Imlog}_2\left(\frac{343}{1136}\right) = 0,883645341$$

$$\text{Diragukan} = \left(-\frac{0}{701}\right) * \text{Imlog}_2\left(\frac{0}{701}\right) + \left(-\frac{701}{701}\right) * \text{Imlog}_2\left(\frac{701}{701}\right) = 0$$

f. Menghitung Gain 1.2

$$\text{Gain (Total, Jumlah Tanggungan)} = 0,986490693 - \left(\frac{1328}{1837} * 0,972599843\right) + \left(\frac{509}{1837} * 0\right) = 0,283380954$$

$$\text{Gain (Total, Total Bulan)} = 0,986490693 - \left(\frac{1136}{1837} * 0,883645341\right) + \left(\frac{701}{1837} * 0\right) = 0,440044799$$

g. Menghitung Entropy 1.3

Node 1.3	Jumlah Kasus (S)	Layak (S1)	Tidak Layak (S2)	Entropy	Gain
Total Bulan (Lancar)	1136	793	343	0,883645341	
Jumlah Tanggungan	Lancar	793	0	0	0,883645341
	Diragukan	343	0	343	

Gambar 5 Node 1.3

$$\text{Entropy (Total Bulan (Lancar))} = \left(-\frac{793}{1136}\right) * \text{Imlog}_2\left(\frac{793}{1136}\right) + \left(-\frac{343}{1136}\right) * \text{Imlog}_2\left(\frac{343}{1136}\right) = 0,883645341$$

**Entropy (Jumlah Tanggungan)**

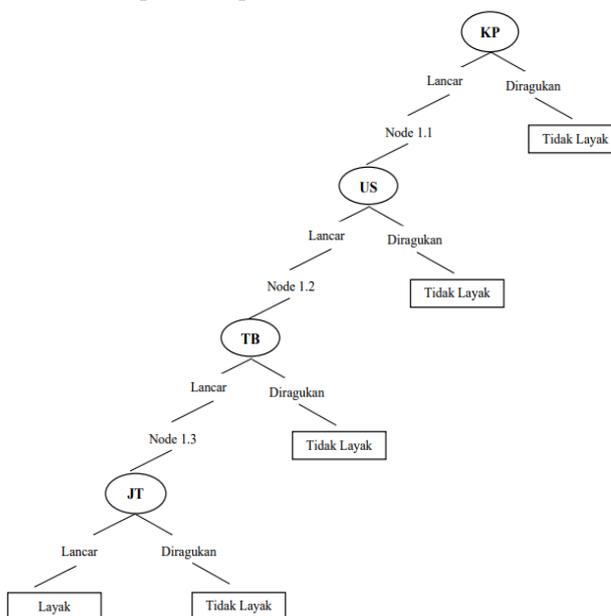
$$\text{Lancar} = \left(-\frac{793}{793}\right) * \text{Imlog}_2\left(\frac{793}{793}\right) + \left(-\frac{0}{793}\right) * \text{Imlog}_2\left(\frac{0}{793}\right) = 0$$

$$\text{Diragukan} = \left(-\frac{0}{343}\right) * \text{Imlog}_2\left(\frac{0}{343}\right) + \left(-\frac{343}{343}\right) * \text{Imlog}_2\left(\frac{343}{343}\right) = 0$$

h. Menghitung Gain 1.3

$$\text{Gain (Total, Jumlah Tanggungan)} = 0,883645341 - \left(\frac{793}{1136} * 0\right) + \left(\frac{343}{1136} * 0\right) = 0,883645341$$

i. Hasil Perhitungan Excel berdasarkan pohon keputusan



Gambar 6 Pohon Keputusan

j. Hasil berdasarkan akurasi menggunakan confusion matrix

Tabel 3 Confusin Matrix Pada Excel

	<b>Layak</b>	<b>Tidak Layak</b>		
<b>Layak</b>	TP	FN	<b>Layak</b>	793
<b>Tidak Layak</b>	FP	TN	<b>Tidak Layak</b>	0
			<b>Total</b>	6076

$$Accuracy = \frac{\text{jumlah prediksi benar}}{\text{jumlah total prediksi}} = \frac{793+5282}{6076} * 100 \% = \mathbf{0,999 \%}$$

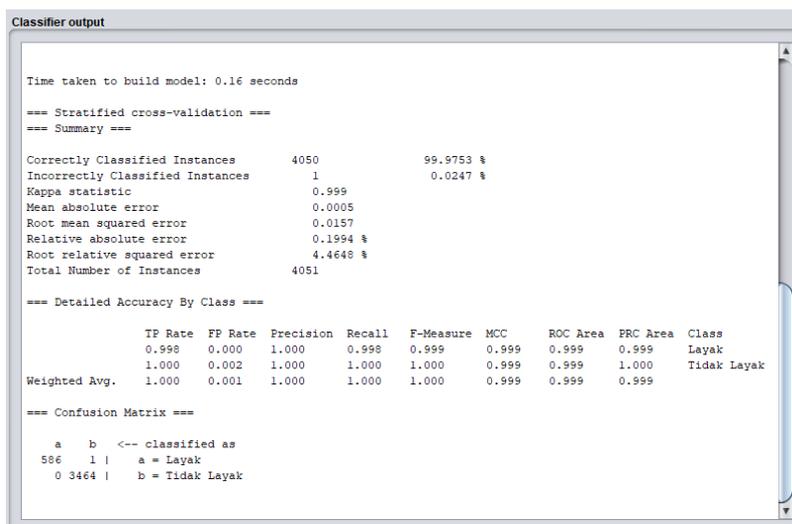
$$Precision = \frac{TP}{TP+FP} = \frac{586}{586+0} * 100 \% = \mathbf{1 \%}$$

$$Recall = \frac{TP}{TP+FN} = \frac{586}{586+1} * 100 \% = \mathbf{0,998 \%}$$

$$F-Measure = 2 * \frac{Precision*Recall}{Precision+Recall} = 2 * \frac{1*0,998}{1+0,998} 100 \% = \mathbf{0,999 \%}$$

2. Hasil Pengujian Algoritma Decision Tree C4.5 Menggunakan Weka

hasil dari pengujian data *testing* presentase yang di prediksi secara benar *correctly classified instance* sebesar 99,9753% dan *incorrectly classified instance* sebesar 0,0247%. Dari 4051 data terdapat 4050 data yang berhasil diklasifikasikan dengan benar dan 1 data yang tidak dapat diklasifikasikan dengan benar.



Gambar 7 Hasil Pengujian Menggunakan Weka

3. Hasil Pengujian Algoritma Menggunakan Python

hasil pengujian algoritma *decision tree c4.5* menggunakan python ini mendapatkan nilai akurasi sebesar 0,9997531473710195.

```

[34] xtrain, xtest, ytrain, ytest = train_test_split(atr_data, cls_data, test_size=0.4, random_state=10127)
tree_data = DecisionTreeClassifier(random_state=10127)
tree_data.fit(xtrain, ytrain)

DecisionTreeClassifier(ccp_alpha=0.0, class_weight=None, criterion='gini',
max_depth=None, max_features=None, max_leaf_nodes=None,
min_impurity_decrease=0.0, min_impurity_split=None,
min_samples_leaf=1, min_samples_split=2,
min_weight_fraction_leaf=0.0, presort='deprecated',
random_state=10127, splitter='best')

[35] print("Nilai Akurasi data:", tree_data.score(xtest, ytest))

Nilai Akurasi data: 0.9997531473710195
    
```

Gambar 8 Hasil Akurasi Pada Python

**B. Pembahasan Penelitian**

Pada penelitian Klasifikasi Permasalahan Kredit Macet Pada Bank Menggunakan Algoritma Decision Tree C4.5 dengan menggunakan tiga pengujian yaitu perhitungan manual menggunakan excel, pengujian menggunakan weka, dan pengujian menggunakan python menghasilkan nilai *accuracy* sebesar 99,9753%, *precision* sebesar 100%, *recall* sebesar 99,8%, dan *f-measure* sebesar 99,9%.

**IV. KESIMPULAN DAN SARAN****A. Kesimpulan**

Penerapan algoritma *decision tree* C4.5 dalam permasalahan kredit macet pada bank diukur berdasarkan nilai *accuracy*, *precision*, *recall*, dan *f-measure* dengan pembagian data 60 : 40 dimana 60% *data training* dan 40% *data testing*. Hasil pengujian dengan menggunakan *tool* weka memiliki nilai *accuracy* sebesar 99,9753%, *precision* sebesar 100%, *recall* sebesar 99,8%, dan *f-measure* sebesar 99,9%.

**B. Saran**

Saran untuk penelitian selanjutnya diharapkan dapat menggunakan algoritma klasifikasi lainnya agar dapat diketahui perbedaan antara algoritma *decision tree* c4.5 dan algoritma klasifikasi lainnya dan sekaligus sebagai referensi bagi penelitian selanjutnya dalam meneliti permasalahan yang sama.

**PENGAKUAN**

Naskah ilmiah ini adalah sebagian dari penelitian Tugas Akhir milik Anida Nirwana dengan judul Klasifikasi Permasalahan Kredit Macet Pada Bank Menggunakan Algoritma *Decision Tree* C4.5, yang dibimbing oleh Amril Mutoi Siregar dan Rahmat.

**DAFTAR PUSTAKA**

- [1] R. Widayati and W. E. Mendari, "Upaya Penanganan Kredit Bermasalah Pada Bank Nagari Cabang Utama Padang," pp. 1–12, 2019, doi: 10.31219/osf.io/ewm65.
- [2] S. D. Kamil, D. Widiyanto, N. Chamidah, P. S. Informatika, F. I. Komputer, and D. Tree, "Perbandingan Metode Decision Tree Dengan Naïve Bayes," *Senamika*, pp. 539–550, 2020.
- [3] S. Wahyuningsih, D. R. Utari, U. B. Luhur, D. Tree, and K. Validation, "Perbandingan Metode K-Nearest Neighbor, Naïve Bayes dan Decision Tree untuk Prediksi Kelayakan Pemberian Kredit," pp. 8–9, 2018.
- [4] F. M. Hana, "Klasifikasi Penderita Penyakit Diabetes Menggunakan Algoritma Decision Tree C4 . 5," 2020.
- [5] E. Karyadiputra and N. Hijriana, "Penerapan Algoritma Decision Tree C4.5 Untuk Klarifikasi Penentuan Daftar Prioritas Pengembangan Jembatan," *Technologia*, vol. 10, no. 1, pp. 43–46, 2019.
- [6] M. P. A. S. R. R. A. P. Pritasari, "Klasifikasi Untuk Memprediksi Pembayaran Kartu Kredit Macet Jurnal Teknologi," vol. 3, no. 1, pp. 91–101, 2020.
- [7] E. Ajeng, A. Putri, E. Nuraina, and E. E. Yusdita, "Upaya Pencegahan dan Penanganan Kredit Macet Ditinjau dari Persepsi Nasabah," vol. 7, no. 2, pp. 185–196, 2020.
- [8] L. N. Rani, "KLASIFIKASI NASABAH MENGGUNAKAN ALGORITMA C4.5 SEBAGAI DASAR PEMBERIAN KREDIT," vol. 2, no. 2, pp. 33–38, 2015.
- [9] W. Darmawan, A. Klasifikasi, D. Tree, N. Bayes, and D. Mining, "ALGORITMA KLASIFIKASI DATA MINING UNTUK PREDIKSI STATUS MAHASISWA MENGGUNAKAN ALGORITMA DECISION TREE (C4.5)," vol. XII, no. 1, pp. 15–22, 2017.
- [10] E. Paskalis, K. Orpa, and E. F. Ripanti, "Model Prediksi Awal Masa Studi Mahasiswa," vol. 7, no. 4, pp. 272–278, 2019.
- [11] S. A. Qutsiah, M. K. Sophan, and Y. F. Hendrawan, "DATAR MENGGUNAKAN PYTHON PADA PERANGKAT," vol. XI, 2016.
- [12] M. Faid, M. Jasri, and T. Rahmawati, "Perbandingan Kinerja Tool Data Mining Weka dan Rapidminer Dalam Algoritma Klasifikasi," *Teknika*, vol. 8, no. 1, pp. 11–16, 2019, doi: 10.34148/teknika.v8i1.95.