

# Analisis Data Mining Menggunakan Metode Algoritma C4.5 Menentukan Penerima Bantuan Langsung Tunai

Lukman Bachtiar<sup>1</sup>, Mahradianur<sup>2\*</sup>

<sup>1,2</sup> Fakultas Ilmu Komputer, Universitas Darwan Ali, Kota Sampit, Indonesia

e-mail: <sup>1</sup>lukman.bachtiar@gmail.com, <sup>2</sup>mahradianur@gmail.com

---

Informasi Artikel      Diterima: 25-01-2023      Direvisi: 24-02-2023      Disetujui: 27-02-2023

---

## Abstrak

Bantuan Langsung Tunai (BLT) adalah bantuan yang dibuat oleh pemerintah untuk orang atau keluarga yang kurang mampu dan rentan akan kemiskinan dengan mereka yang memenuhi persyaratan yang ada pada program. Tujuan penelitian ini untuk memberikan masukan tentang menentukan apakah keluarga tersebut masih berhak menerima atau tidak berhak menerima bantuan BLT, dimana masih ada keluarga yang berhak menerima dan memenuhi syarat dari program bantuan pemerintahan ini. Sumber data yang diperoleh dapat dari Kantor Desa Regei Lestari. Metode yang dipakai pada penelitian merupakan teknik data mining dengan algoritma C4.5 yang diterapkan pada software RapidMiner. Atribut yang dipakai untuk menentukan keluarga yang berhak menerima dan tidak berhak menerima pada program bantuan BLT yaitu Penghasilan, Status, Pekerjaan, Umur. Hasil dari klasifikasi data menggunakan metode algoritma C4.5 dan pengujian dengan RapidMiner adalah Jika Pekerjaan = buruh tani maka hasilnya adalah menerima dengan nilai 16 item dari menerima dan 0 yang tidak menerima, maka didapatkan faktor yang menentukan yang berhak menerima dan yang tidak berhak menerima bantuan langsung tunai (BLT).

Kata Kunci: Algoritma C45, Data Mining, BLT

## Abstract

*Direct cash assistance (BLT) is assistance provided by the government for underprivileged and vulnerable people or families with those who meet the requirements of the program. The purpose of this research is to provide input on determining whether the family is still entitled to receive or not entitled to receive BLT assistance, where there are still families who are entitled to receive and meet the requirements of this government assistance program. The source of the data obtained can be from the Regei Lestari Village Office. The method used in this study is a data mining technique with the C4.5 algorithm applied to the RapidMiner software. The attributes used to determine families who are entitled to receive and who are not entitled to receive the BLT assistance program are Income, Status, Occupation, Age. The results of data classification using the C4.5 algorithm method and testing with RapidMiner are If Occupation = farmhand then the result is accepting with a value of 16 items from accepting and 0 not accepting, then the determining factor is obtained who is entitled to receive and who is not entitled to receive assistance direct cash (BLT).*

Keywords: C45 Algorithm, Data Mining, BLT

## 1. Pendahuluan

Data mining merupakan proses menemukan korelasi baru yang bermanfaat, pola dan *trend* dengan menambang sejumlah *repository* data dalam jumlah besar, menggunakan teknologi pengenalan pola seperti statistik dan teknik matematika (Fatmawati and Windarto 2018). Teknik-teknik, metode-metode, atau algoritma dalam data mining sangat bervariasi. Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses *Knowledge Discovery in*

*Database* (KDD) secara keseluruhan (Yuli Mardi 2019). Data mining adalah proses yang mempekerjakan satu atau lebih teknik pembelajaran komputer (*machine learning*) untuk menganalisis dan mengekstraksi pengetahuan (*knowledge*) secara otomatis (Anisa and Mesran 2018).

Klasifikasi merupakan salah satu teknik dalam data mining. Klasifikasi (taksonomi) merupakan proses penempatan objek atau konsep tertentu ke dalam satu set kategori berdasarkan objek yang digunakan. Salah satu



teknik klasifikasi yang populer digunakan adalah *decision tree*. (Ardiansyah and Walim 2018). Klasifikasi juga dapat diartikan sebagai sebuah proses menemukan suatu model atau fungsi yang menggambarkan dan membedakan kelas data atau konsep dengan tujuan dapat menggunakan model untuk membuat prediksi kelas objek dimana kelas labelnya tidak diketahui (Suparyanto dan Rosad 2020). Tujuan dari klasifikasi adalah untuk menemukan model dari training set yang membedakan atribut ke dalam kategori atau kelas yang sesuai, model tersebut kemudian digunakan untuk mengklasifikasikan atribut yang kelasnya belum diketahui sebelumnya (Azwanti 2018).

Algoritma C4.5 adalah program yang memberi kontribusi satu set data berlabel dan menghasilkan pohon keputusan sebagai keluaran. Pohon keputusan tindak lanjut ini kemudian diverifikasi terhadap data uji berlabel yang tidak terlihat untuk menghitung generalisasinya (Sugara, Adidarma, and Budilaksono 2019). Algoritma C4.5 merupakan salah satu algoritma yang digunakan untuk melakukan klasifikasi atau segmentasi yang bersifat *prediktif* (P et al. 2019).

Pohon keputusan merupakan metode klasifikasi dan prediksi yang sangat kuat dan terkenal. Dimana metode pohon keputusan mengubah fakta yang sangat besar menjadi pohon keputusan yang dapat mempresentasikan aturan-aturan. Aturan tersebut dapat dengan mudah dipahami ataupun dijelaskan dengan ulang karena memiliki bahasa alami yang dapat dimengerti (Ucha Putri et al. 2021). *Decision tree* atau pohon keputusan menggambarkan urutan struktur seperti pohon (*tree*) dimana setiap node merepresentasikan atribut, cabangnya merepresentasikan nilai dari atribut dan daun merepresentasikan kelas. Node yang paling atas dari pohon keputusan disebut sebagai root (Ali Khumaidi, Ridwan Raafi'udin 2020). Pohon Keputusan Juga berguna untuk mengeksplorasi data, menemukan hubungan tersembunyi (Faisal 2019).

Bantuan langsung tunai (BLT) adalah program pemerintah kepada keluarga ataupun individu yang dinilai kurang mampu atau rentan akan kemiskinan. Bantuan yang diberikan berupa sejumlah uang tunai kepada keluarga ataupun individu yang dinilai kurang mampu. Dalam pelaksanaannya, program BLT secara langsung berdampak terhadap peningkatan kemampuan beli masyarakat miskin dan berdampak juga kepada pengelolaan dana yang digunakan untuk mencukupi kebutuhan minimal setiap hari (Dewi and Andrianus 2021). Pada penelitian ini masyarakat yang dinilai kurang mampu dan memerlukan bantuan adalah masyarakat yang kurang

mampu dan terkena dampak pengaruh dari wabah pandemi Covid-19 di masa itu.

Adapun dalam penulisan penelitian ini berlandaskan pada permasalahan yang terjadi Pada desa Regei Lestari Kecamatan Teluk Sampit Kabupaten Kotawaringin Timur Provinsi Kalimantan Tengah pada tahun 2021. Desa Regei Lestari tiap tahun nya selalu melaksanakan program Bantuan Langsung Tunai kepada warga desanya. Pelaksanaan program BLT di desa Regei Lestari di peruntukan kepada warga masyarakat kurang mampu dan terkena dampak dari Musibah COVID 19 sehingga dari terkenanya dampak Dari Virus COVID 19 maka bertambah juga jumlah warga masyarakat yang memerlukan bantuan pemerintah ini, hingga pada akhirnya mengakibatkan perangkat desa kesulitan dalam menentukan keluarga atau masyarakat yang berhak menerima bantuan tersebut.

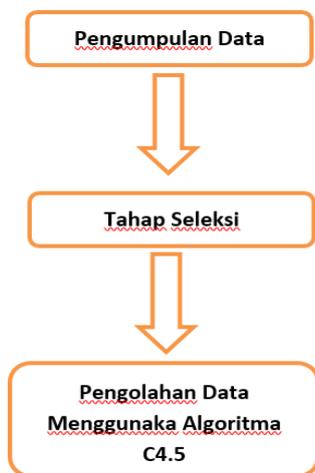
Setelah melakukan pencarian dan pengumpulan data penulis melihat beberapa kekurangan pada proses pelaksanaan pemberian bantuan langsung tunai (BLT) karena proses menentukannya masih dilakukan secara manual yang berdampak kurang tepat sasaran pada orang yang berhak menerima bantuan tersebut. Sehingga menimbulkan masyarakat yang penerima bantuan masih tergolong mampu yang menerima bantuan tersebut. Oleh karena dengan Metode Algoritma C4.5 diharapkan dapat membantu perangkat Desa untuk mencari masyarakat yang berhak menerima bantuan langsung tunai (BLT).

Pada penelitian yang berjudul "Analisis Data Mining Menentukan Penerima Bantuan Langsung Tunai pada Desa Pamatang Purba dengan Algoritma C 4.5" (Sinaga, Winanjaya, and Susianti 2021) memiliki beberapa perbedaan pada penelitian yang dilakukan sekarang, diantaranya tempat objek penelitian yang berbeda dikarena penulis melakukan penelitian pada Desa Regei Lestari, dan memiliki jumlah data yang lebih banyak dari pada penelitian yang diatas. Metode dalam penelitian ini menggunakan metode survey lapangan dengan memberikan batasan kriteria pada pendapatan bulanan, jumlah tanggungan, jenis tempat tinggal, dan kendaraan (Junaidi, Yunita, et all, 2023)

Pada penelitian ini bertujuan untuk membantu Perangkat Kantor desa Regei Lestari untuk mencari masyarakat yang berhak menerima bantuan langsung tunai (BLT) secara efektif dan tepat sasaran. Penelitian ini juga bermanfaat memberikan rekomendasi untuk perangkat Kantor Desa Regei Lestari dalam menentukan calon penerima BLT, Supaya untuk kedepannya Perangkat dari Kantor desa Regei Lestari membuat keputusan secara tepat

sasaran kepada masyarakat yang menerima bantuan langsung tunai (BLT).

## 2. Metode Penelitian



Gambar 1 Alur Penelitian

### 1. Pengumpulan data

Data yang digunakan dalam penelitian adalah data penerima Bantuan Langsung Tunai (BLT) tahun 2020/2021 pada Desa Regei Lestari yang Diambil dari kantor Desa Regei Lestari yang mengadakan bantuan langsung tunai tersebut. Jumlah data yang digunakan ada 121 data yang terdiri dari RT 01 sampai RT 08 di desa tersebut.

### 2. Tahap Seleksi

Tahap penyeleksian data perlu untuk dilakukan sebelum tahap pemrosesan data. Pada tahap penyeleksian data, data yang digunakan tadi diseleksi untuk mencari data atribut yang penting dalam penelitian. Data yang telah diseleksi menggunakan metode manual di MS. Excel berupa, penghasilan dengan *value* 300, 400, 600, 700, 800, 900, umur dengan *value* dewasa dan muda, status dengan *value* menikah, janda dan duda, pekerjaan dengan *value* ibu rumah tangga, buruh tani dan petani. Yang semua berjumlah 5 atribut.

### 3. Proses Pengolahan Data Dengan Algoritma C.5

Pada tahap ini data yang telah diseleksi sesuai dengan keperluan penelitian kemudian di proses dengan cara menghitung manual pada MS. Excel menggunakan rumus – rumus Algoritma C4.5. Adapun rumus – rumus yang digunakan dalam perhitungan pengolahan data dengan algoritma sebagai berikut:

#### a) Entropy

Entropy (S) merupakan jumlah bit yang diperkirakan dibutuhkan untuk dapat

mengekstrak suatu kelas (+ atau -) dari sejumlah data acak pada ruang sampel S. Entropy dapat dikatakan sebagai kebutuhan bit untuk menyatakan suatu kelas. Entropy digunakan untuk mengukur ketidakkaslian S. Untuk perhitungan nilai Entropy sebagai berikut:

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

Keterangan:

S = Himpunan kasus

n = Jumlah partisi S

$p_i$  = Proporsi  $S_i$  terhadap S

b) Gain

Gain (S,A) merupakan perolehan informasi dari atribut A relative terhadap output data S. Perolehan informasi didapat dari output data atau variable dependent S yang dikelompokkan berdasarkan atribut A, dinotasikan dengan gain (S,A). Untuk memilih atribut sebagai akan didasarkan pada nilai gain tertinggi dari atribut-atribut yang ada. Dengan rumus sebagai berikut:

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

Keterangan:

S = Himpunan kasus

A = Atribut

n = Jumlah partisi atribut A

$|S_i|$  = Proporsi  $S_i$  terhadap S

$|S|$  = jumlah kasus dalam S

#### c) Split Info

Split Info adalah rumus yang digunakan untuk memisahkan sejumlah atribut data yang telah dihasilkan dari dari nilai gain untuk perhitungannya sebagai berikut:

$$\text{SplitInfo} = \sum_{i=1}^k p(v_i|s) \log_2 p(v_i|s) \quad (3)$$

Keterangan:

S = Jumlah data

$S_i$  = Jumlah masing-masing pada setiap atribut.

#### d) Gain rasio

Gain rasio adalah rumus yang digunakan untuk menentukan dimana perhitungan selanjutnya yang dihitung ulang lagi dan rumusnya sebagai berikut:

$$\text{Gain Ratio (s, j)} = \frac{\text{Gain(S, j)}}{\text{SplitInfo(S, j)}} \quad (4)$$

Setelah mendapatkan nilai dari perhitungan tersebut, maka bisa membuat pohon keputusan dari perhitungan tersebut. pohon keputusan dapat diartikan suatu cara untuk memprediksi atau mengklarifikasi yang sangat kuat. Pohon keputusan dapat membagi kumpulan data yang besar menjadi himpunan-himpunan record yang lebih kecil dengan menerapkan serangkaian aturan keputusan (Batubara and Windarto 2019). Pohon keputusan adalah hasil dari proses perhitungan

entropy dan information gain, setelah perhitungan berulang-ulang sampai semua atribut pohon memiliki kelas dan tidak bisa lagi dilakukan proses perhitungan (Cynthia and Ismanto 2018).

**3. Hasil dan Pembahasan**

Penelitian ini dilakukan pada desa Regei Lestari. Data penelitian yang digunakan dari penulisan ini adalah data yang ada pada kantor desa Regei Lestari dalam proses pengumpulan data penelitian. Berikut adalah data yang dipakai :

Tabel 1. Tabel data penerima bantuan langsung tunai 2020/2021

Nama	Penghasilan	status	Pekerjaan	Umur	Penerima BLT
SUMIN AH	500,000	Menikah	Ibu Rumah Tangga	38	Tidak Menerima
SARIF UDIN SUSA NTI	700,000 300,000	Duda Menikah	Buruh Tani Ibu Rumah Tangga	40 36	Menerima Tidak Menerima
.....	.....	...	.....	...	...
JUKIR LI	500,000	Menikah	Petani	43	Menerima
PARID AH	400,000	Janda	Ibu Rumah Tangga	34	Tidak Menerima

Dari tabel 1 diatas dalam menentukan hasil untuk mencari model aturan pertama yang menggunakan Data Mining dengan Algoritma C4.5. Algoritma C4.5 dilakukan untuk memberikan model aturan dari data yang digunakan. Berikut ini langkah-langkah dalam membentuk pohon keputusan dengan menggunakan Algoritma C4.5 :

Dari table data diatas diketahui kasusnya ada 121 record, dan di ketahui label menerima ada 40 dan tidak menerima ada 81 record sehingga Entropy nilai akar total yang didapat:

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

$$= (-46/121 . \log_2 (46/121)) + (-75/121 . \log_2 (75/121))$$

$$= 0.576627877$$

Menghitung nilai Entropy pada atribut Penghasilan dengan kategori 300, jumlah kasus terdapat 40 record, terdapat label menerima sebanyak 15 record dan label tidak menerima sebanyak 25 record.

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

$$= (-17/40 . \log_2 (17/40)) + (-23/40 . \log_2 (23/40))$$

$$= 0.588343259$$

Seterusnya perhitungan Entropy pada atribut Penghasilan sampai dengan kategori 900. Selanjutnya menghitung nilai Gain pada Atribut Penghasilan dengan kategori 300 terdapat 40 record, kategori 400 terdapat 17 record, kategori 500 terdapat 22 record, kategori 700 terdapat 37 record, kategori 800 terdapat 3 record, dan kategori 900 terdapat 2 record :

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

$$= (21)*0.458336042 + ((22/121)*0.594841487) + ((37/121)*0.582574061) + ((3/121)*0.577464587) + ((2/121)*0.301029996)$$

$$= -0.012151432$$

Selanjutnya Menghitung SplitInfo pada atribut Penghasilan berdasarkan jumlah kasus pada Penghasilan:

$$SplitInfo = \sum_{i=1}^k p(v_i|S) \log_2 p(v_i|S) \quad (3)$$

$$= -((40/121.\log_2 (40/121)) + (17/121.\log_2 (17/121)) + (22/121.\log_2 (22/121)) + (37/121.\log_2 (37/121)) + (3/121.\log_2 (3/121)) + (2/121.\log_2 (2/121)))$$

$$= -3.102589432$$

Kemudian menghitung Nilai GainRatio Penghasilan sebagai nilai yang digunakan sebagai pemecahan Node pada pohon keputusan.

$$Gain Ratio (s, j) = \frac{Gain(S, j)}{SplitInfo(S, j)} \quad (4)$$

$$= -0.012151432 / -3.102589432$$

$$= 0.003916545$$

Hasil dari perhitungan nilai Entropy, nilai Gain, SplitInfo dan GainRatio untuk tiap atribut dapat dilihat pada Tabel diatas sebagai berikut :

Tabel 2. Tabel Perhitungan gain node 1

node	atribut	value	jumlah kasus	menerima	tidak menerima	entropi	gain	split info	gainrasio
1	total penghasilan		121	40	81	-0.57663			
		300	40	15	25	-0.58834	-0.01215	-3.10259	0.003917
		400	17	2	15	-0.45834			
		500	22	9	13	-0.59484			
		700	37	13	24	-0.58257			
		800	3	1	2	-0.57746			
	900	2	0	2	-0.30103				
	umur	dewa					-0.0623	-1.05692	0.058946
		sa tua	80	12	68	-0.48461			
	status		41	28	13	-0.57231	-0.09702	-1.37493	0.070561
		menikah	77	24	53	-0.57049			
		janda	34	6	28	-0.50341			
	perkerjaan	duda	10	10	0	-0.30103	-0.23215	-1.05348	0.220369
		ibu rumah tangga	91	10	81	-0.45142			
		buruh tani	16	16	0	-0.30103			
		petani	14	14	0	-0.30103			

Dari Tabel 2 diketahui nilai GainRatio terbesar adalah atribut pekerjaan dengan nilai 0.220369252 dengan kategori ibu rumah tangga dengan nilai -0.451419888, kategori buruh tani dengan nilai -0.301029996, dan kategori petani dengan nilai -0.301029996 dan dijadikan simpul akar untuk mengulangi perhitungan nilai Entropy dan nilai Gain selanjutnya. Hasil perhitungan nilai Entropy dan nilai Gain untuk atribut Pekerjaan untuk kategori Ibu rumah tangga, dapat dilihat pada Tabel 3. Sebagai berikut:

Tabel 3 Dari Tabel 2 diketahui nilai GainRatio terbesar adalah atribut pekerjaan dengan nilai 0.220369252 dengan kategori ibu rumah tangga dengan nilai -0.451419888, kategori buruh tani dengan nilai -0.301029996, dan kategori petani dengan nilai -0.301029996 dan dijadikan simpul akar untuk mengulangi perhitungan nilai Entropy dan nilai Gain selanjutnya. Hasil perhitungan nilai Entropy dan nilai Gain untuk atribut pekerjaan untuk kategori ibu rumah tangga, dapat dilihat pada Tabel 3. Sebagai berikut:

Tabel 3. Perhitungan gain node 2

node	atribut	value	jumlah kasus	menerima	tidak menerima	entropi	gain	split info	gainrasio	
2	total penghasilan		91	10	81	0.45142				
		300	28	3	25	0.448907	0.0283	2.297844	0.012316	
		400	15	0	15	0.30103				
		500	19	6	13	0.57188				
		700	25	1	24	0.373967				
		800	2	0	2	0.30103				
	900	2	0	2	0.30103					
	umur	dewasa		68	0	68	0.30103	0.075242	0.899385	0.083659
		tua	23	10	13	0.598355				
	status			57	4	53	0.411381	0.005654	0.914792	0.006181
		menikah	34	6	28	0.503411				
		janda	0	0	0	0				

Dari Tabel 3 diketahui nilai GainRatio terbesar adalah atribut umur dengan nilai 0.083659158 dengan kategori dewasa dengan nilai 0.301029996 dan kategori tua dengan nilai 0.598355067 dan dijadikan simpul akar untuk

mengulangi perhitungan nilai Entropy dan nilai Gain selanjutnya pada atribut umur dengan kategori tua, sebagai berikut:

Tabel 4. Perhitungan gain node 3

node	atribut	value	jumlah kasus	tidak menerima	menerima	entropi	gain	split info	gainrasio
3	total penghasilan		24	13	11	0.60055			
							0.017182	2.025804	0.008482
		300	10	5	5	0.60206			
		400	2	2	0	0.30103			
		500	9	3	6	0.577465			
		700	4	3	1	0.545249			
		800	0	0	0	0			
	900	0	0	0	0				
	status						0.023844	1.154797	0.020648
		menikah	13	9	4	0.569095			
		janda	11	4	7	0.585702			
		duda	0	0	0	0			

Dari Tabel 4 gain node 3 diketahui nilai GainRatio terbesar adalah atribut status dengan nilai 0.020647657 dengan kategori menikah dengan nilai 0.56909469, kategori janda dengan nilai 0.585702113, dan kategori duda bernilai 0. Dan dijadikan simpul akar untuk mengulangi

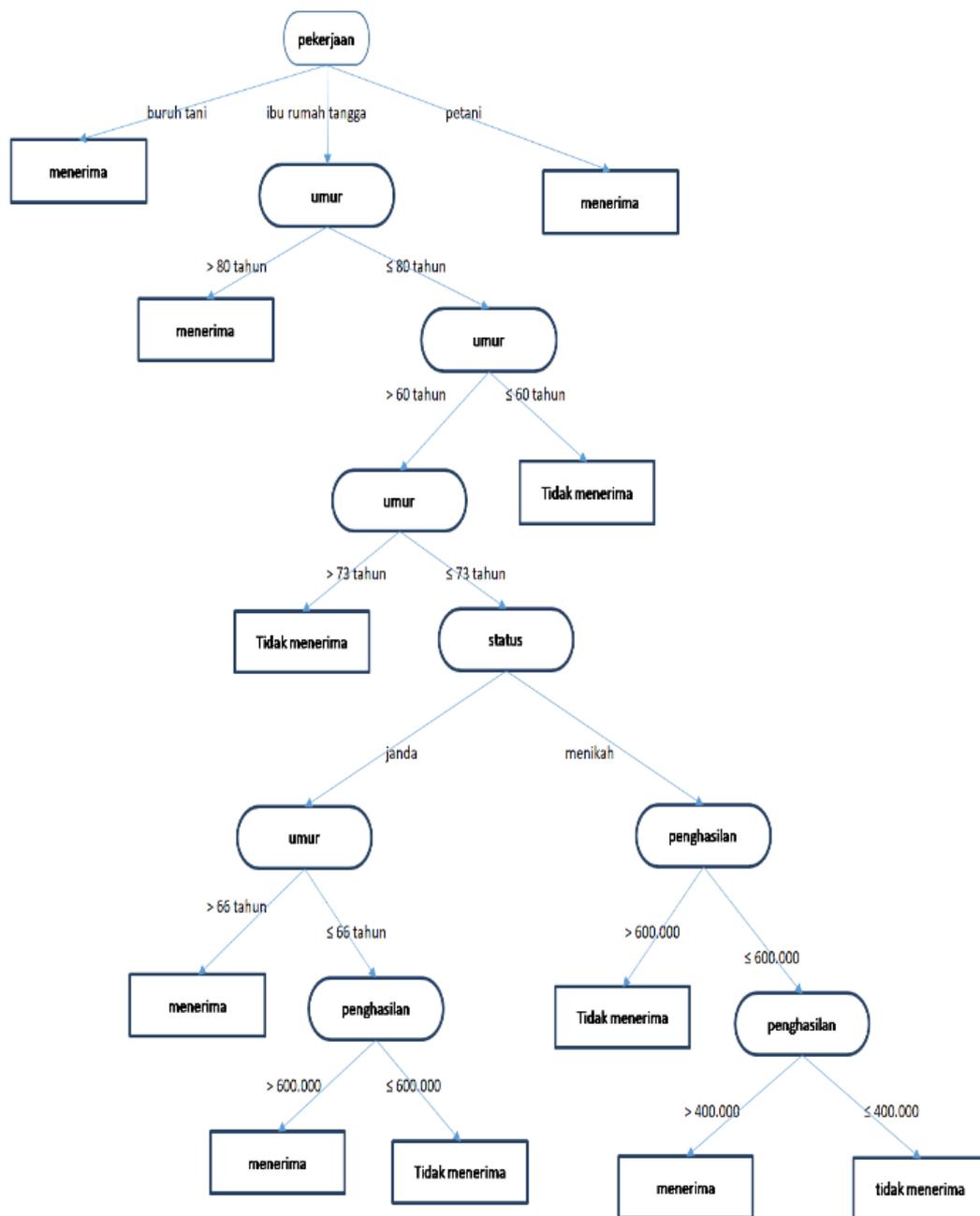
perhitungan nilai Entropy dan nilai Gain selanjutnya, untuk mencari akar berikutnya pada atribut status kategori menikah pada gain node 4 sebagai berikut:

Tabel 5. Perhitungan gain node 4

node	atribut	value	jumlah kasus	menerima	tidak menerima	entropi	gain	split info	gainrasio
4	total penghasilan		12	3	9	0.545249			
							0.053325	1.722472	0.030958
		300	5	1	4	0.518352			
		400	1	0	1	0.30103			
		500	4	2	2	0.60206			
		700	2	0	2	0.30103			
		800	0	0	0	0			
		900	0	0	0	0			

Dari Tabel 5 diketahui nilai GainRatio terbesar adalah atribut penghasilan dengan nilai 0.030958 dengan kategori 300 dengan nilai 0.518352007, kategori 400 dengan nilai 0.301029996, kategori 500 dengan nilai 0.602059991, kategori 700 dengan nilai

0.301029996 dan kategori 800, 900 dengan nilai 0. Setelah tidak ada proses dilakukan lagi, diperoleh pohon keputusan yang dapat dilihat pada Gambar 2 berikut :



Gambar 2. Pohon keputusan

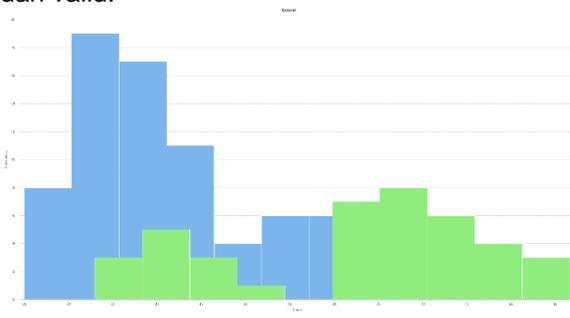
Berdasarkan bentuk pohon keputusan pada gambar 2, aturan atau Rule model yang didapatkan adalah sebagai berikut:

- 1) Jika Pekerjaan = buruh tani maka hasilnya = Menerima {Tidak Menerima=0, Menerima=16}
- 2) Jika Pekerjaan = Ibu Rumah Tangga, dan umur = > 80 tahun maka hasilnya = Menerima {Tidak Menerima=0, Menerima=1}
- 3) Jika pekerjaan = ibu rumah tangga, umur = ≤ 80 sampai umur ≤ 60 maka hasilnya = Tidak Menerima {Tidak Menerima=68, Menerima=0}
- 4) Jika pekerjaan = ibu rumah tangga, umur = ≤ 80 dan > 60, > 73 maka hasilnya = Tidak Menerima {Tidak Menerima=5, Menerima=0}
- 5) Jika pekerjaan = ibu rumah tangga, umur = ≤ 80 dan > 60, ≤ 73 dengan status = janda yang berumur > 66 maka hasilnya = Menerima {Tidak Menerima=0, Menerima=4}
- 6) Jika pekerjaan = ibu rumah tangga, umur = ≤ 80 dan > 60, ≤ 73 dengan status = janda yang berumur ≤ 66 tahun dan penghasilan = > 600.000 maka hasilnya = Menerima {Tidak Menerima=0, Menerima=1}

- 7) Jika pekerjaan = ibu rumah tangga, umur =  $\leq 80$  dan  $> 60, \leq 73$  dengan status = janda yang berumur  $\leq 66$  tahun dan penghasilan =  $\leq 600.000$  maka hasilnya = Tidak Menerima {Tidak Menerima=1, Menerima=0}
- 8) Jika pekerjaan = ibu rumah tangga, umur =  $\leq 80$  dan  $> 60, \leq 73$  dengan status = menikah dan penghasilan  $> 600.000$  maka hasilnya = Tidak Menerima {Tidak Menerima=2, Menerima=0}
- 9) Jika pekerjaan = ibu rumah tangga, umur =  $\leq 80$  dan  $> 60, \leq 73$  dengan status = menikah dan penghasilan  $\leq 600.000$  Sampai  $> 400.000$  maka hasilnya = Menerima {Tidak Menerima=1, Menerima=3}
- 10) Jika pekerjaan = ibu rumah tangga, umur =  $\leq 80$  dan  $> 60, \leq 73$  dengan status = menikah dan penghasilan  $\leq 600.000$  Sampai  $\leq 400.000$  maka hasilnya = Tidak Menerima {Tidak Menerima=4, Menerima=1}
- 11) Jika pekerjaan = petani maka hasilnya = Menerima {Tidak Menerima=0, Menerima=14}

Hasil yang dilakukan penulis dalam perhitungan Algoritma C4.5 menghasilkan model aturan dengan pekerjaan tertinggi adalah Jika Pekerjaan = buruh tani maka hasilnya adalah menerima dengan nilai 16 item dari menerima dan 0 yang tidak menerima.

Hasil yang diberikan oleh tools RapidMiner Algoritma C4.5 menghasilkan model aturan dengan pekerjaan tertinggi adalah Jika Pekerjaan = buruh tani maka hasilnya adalah menerima dengan nilai 16 item dari menerima dan 0 yang tidak menerima. Artinya Hasil dari proses yang dilakukan penulis dan tools RapidMiner untuk Algoritma C4.5 adalah sama dan valid.



Gambar 3. Tampilan grafik Pohon Keputusan Rapidminer

#### 4. Kesimpulan

Berdasarkan penelitian yang dilakukan dapat disimpulkan bahwa data mining yang digunakan dengan metode algoritma C45 dapat diterapkan. Dengan penelitian ini hasil algoritma

yang didapatkan dari data yang diperoleh, dapat memberikan rekomendasi untuk pengambil keputusan sehingga Kepala Desa Regei lestari dapat memberikan bantuan BLT Secara merata berdasarkan model yang diberikan.

#### Referensi

- Ali Khumaidi, Ridwan Raafi'udin, Indra Permana Solihin. 2020. "Analisa Kepuasan Konsumen Menggunakan Algoritma C4.5." *Prosiding Seminar Nasional Ilmu Sosial Dan Teknologi* (3):126–31.
- Anisa, Anisa, and Mesran Mesran. 2018. "Analisa Pola Pekerjaan Lulusan Stmik Budi Darma Menerapkan Metode C4.5." *KOMIK (Konferensi Nasional Teknologi Informasi Dan Komputer)* 2(1):446–49. doi: 10.30865/komik.v2i1.974.
- Ardiansyah, Dian, and Walim Walim. 2018. "Algoritma C4.5 Untuk Klasifikasi Calon Peserta Lomba Cerdas Cermat Siswa Smp Dengan Menggunakan Aplikasi Rapid Miner." *Jurnal Inkofar* 1(2):5–12.
- Azwanti, Nurul. 2018. "Analisa Algoritma C4.5 Untuk Memprediksi Penjualan Motor Pada Pt. Capella Dinamik Nusantara Cabang Muka Kuning." *Informatika Mulawarman: Jurnal Ilmiah Ilmu Komputer* 13(1):33. doi: 10.30872/jim.v13i1.629.
- Batubara, Dinda Nabila, and Agus Perdana Windarto. 2019. "Analisa Klasifikasi Data Mining Pada Tingkat Kepuasan Pengunjung Taman Hewan Pematang Siantar Dengan Algoritma." *KOMIK (Konferensi Nasional Teknologi Informasi Dan Komputer)* 3(1):588–92. doi: 10.30865/komik.v3i1.1664.
- Cynthia, Eka Pandu, and Edi Ismanto. 2018. "Metode Decision Tree Algoritma C.45 Dalam Mengklasifikasi Data Penjualan Bisnis Gerai Makanan Cepat Saji." *Jurasik (Jurnal Riset Sistem Informasi Dan Teknik Informatika)* 3(July):1. doi: 10.30645/jurasik.v3i0.60.
- Dewi, Ratna, and Habib Furqony Andrianus. 2021. "Analisis Pengaruh Kebijakan Bantuan Langsung Tunai (BLT) Terhadap Kemiskinan Di Indonesia Periode 2005-2015." *MENARA: Ilmu* 15(2):77–84.
- Faisal, Sutan. 2019. "Klasifikasi Data Mining Menggunakan Algoritma C4.5 Terhadap Kepuasan Pelanggan Sewa Kamera Cikarang." *Techno Xplore: Jurnal Ilmu Komputer Dan Teknologi Informasi* 4(1):1–8. doi: 10.36805/technoxplore.v4i1.541.
- Fatmawati, Kiki, and Agus Perdana Windarto. 2018. "Data Mining: Penerapan Rapidminer Dengan K-Means Cluster Pada Daerah Terjangkit Demam Berdarah Dengue (Dbd) Berdasarkan Provinsi."

- Computer Engineering, Science and System Journal* 3(2):173. doi: 10.24114/cess.v3i2.9661.
- Junaidi, A., Yunita, Y., Agustyani, S., Agustyaningrum, C. I., & Arifin, Y. T. (2023). Klasifikasi Penerima Bantuan Sosial Menggunakan Algoritma C 4.5. *Jurnal Teknik Komputer*, 9(1), 77-82.
- P, Dini Rizky Sitorus, Agus Perdana Windarto, Dedy Hartama, and Irfan Sudahri Damanik. 2019. "Penerapan Klasifikasi C4.5 Dalam Meningkatkan Sistem Pembelajaran Mahasiswa." *KOMIK (Konferensi Nasional Teknologi Informasi Dan Komputer)* 3(1):593–97. doi: 10.30865/komik.v3i1.1665.
- Sinaga, Rajes Wasimson, Riki Winanjaya, and S. Susianti. 2021. "Analisis Data Mining Menentukan Penerima Bantuan Langsung Tunai Pada Desa Pematang Purba Dengan Algoritma C 4.5." *Brahmana : Jurnal Penerapan Kecerdasan Buatan* 3(1):1–9. doi: 10.30645/brahmana.v3i1.86.
- Sugara, Bayu, Dedi Adidarma, and Sularso Budilaksono. 2019. "Perbandingan Akurasi Algoritma C4.5 Dan Naïve Bayes Untuk Deteksi Dini Gangguan Autisme Pada Anak." *Jurnal IKRA-ITH Informatika* 3(1):119–28.
- Suparyanto dan Rosad. 2020. "Penerapan Algoritma C4.5 Pada Program Klasifikasi Mahasiswa Dropout." *Suparyanto Dan Rosad 2020* 5(3):248–53.
- Ucha Putri, Sanni, Eka Irawan, Fitri Rizky, Stikom Tunas Bangsa, Pematangsiantar A.-Indonesia Jln Sudirman Blok No, and Sumatera Utara. 2021. "Implementasi Data Mining Untuk Prediksi Penyakit Diabetes Dengan Algoritma C4.5." *Januari* 2(1):39–46.
- Yuli Mardi. 2019. "Data Mining : Klasifikasi Menggunakan Algoritma C4 . 5 Data Mining Merupakan Bagian Dari Tahapan Proses Knowledge Discovery in Database ( KDD ) . *Jurnal Edik Informatika.*" *Jurnal Edik Informatika* 2(2):213–19.