

## Analisis Data Mining Menggunakan Algoritma C 4.5 Dalam Memprediksi Penerima Bantuan Sosial

Leonardo Yemi<sup>1\*</sup>, Sarjon Defit<sup>2</sup>, Sumijan<sup>3</sup>

<sup>1,2,3</sup>Fakultas Ilmu Komputer, Universitas Putra Indonesia "YPTK" Padang, Indonesia

E-mail: Leonardoyemi28@gmail.com<sup>1</sup>, sarjonnd@yahoo.co.uk<sup>2</sup>, sumijan@upiptk.ac.id<sup>3</sup>

### Abstract

Poverty is one of the highest problems most often experienced by various developing countries, there are many ways to overcome the purpose of social assistance is to overcome poverty, social assistance is usually provided by the government and non-profit organizations to groups of people who have economic limitations. The purpose of this study was to help recipients of social assistance to be right on target and can help people with limitations. One of the techniques used in data analysis is data mining. This study identifies recipients of social assistance using data mining efficiently and fairly. and testing the rapidminer application in the prediction process using the C4.5 algorithm. This research dataset uses 80 data based on data on recipients of social assistance in the Jati sub-district, Padang city. The results of the C4.5 algorithm performance test were able to present prediction analysis output with a very good level of accuracy, namely 93.75%. These results are quite evident that the C4.5 algorithm is able to present maximum prediction output in determining recipients of social assistance in the Jati sub-district, Padang city for the next period. Based on these results, it can facilitate and accelerate decision-making related to determining the receipt of social assistance by applying the C4.5 algorithm and can provide more accurate results.

**Keywords:** Social Assistance, Data Mining, Classification, Rapidminer, C4.5 Algorithm.

### Abstrak

Kemiskinan merupakan salah satu masalah tertinggi yang paling sering dialami oleh berbagai negara berkembang, ada banyak cara untuk menanggulangi Tujuan dari bantuan sosial adalah untuk menuntaskan kemiskinan, bantuan sosial biasanya diberikan oleh pemerintah dan organisasi nirlaba kepada kelompok masyarakat yang memiliki keterbatasan ekonomi. tujuan penelitian ini dilakukan dapat membantu penerima bantuan sosial agar tepat sasaran dan dapat membantu masyarakat dengan keterbatasan salah satu teknik yang digunakan dalam analisis data adalah data mining. Penelitian ini mengidentifikasi penerima bantuan sosial menggunakan data mining secara, efisien, dan adil. serta menguji aplikasi rapidminer dalam proses prediksi menggunakan algoritma C4.5. dataset penelitian ini menggunakan 80 data berdasarkan data penerima bantuan sosial dikelurahan jati kota padang. adapun hasil pengujian kinerja algoritma C4.5 mampu menyajikan keluaran analisis prediksi dengan tingkat akurasi yang sangat baik yaitu sebesar 93.75%. hasil tersebut cukup terbukti bahwa algoritma C4.5 mampu menyajikan keluaran prediksi yang maksimal dalam menentukan penerima bantuan sosial dikelurahan jati kota padang untuk periode selanjutnya. berdasarkan hasil tersebut dapat memperbudah dan mempercepat dalam pengambilan keputusan terkait menentukan penerimaan bantuan sosial dengan menerapkan algoritma C4.5 serta dapat memberikan hasil yang lebih akurat.

**Kata kunci:** Bantuan Sosial, Data Mining, Klasifikasi, Rapidminer, Algoritma C4.5.

## 1. Pendahuluan

*Knowledge Discovery in Database (KDD)* adalah proses yang berfokus pada penggalian atau pencarian informasi yang berguna dari kumpulan data yang besar, prosesnya melibatkan pembersihan data, pengambilan sampel, penskalaan, pengelompokan dan lainnya untuk mendapatkan wawasan yang bermakna [1]. Data mining merupakan sebuah inti dari proses *Knowledge discovery in database (KDD)*, meliputi dugaan algoritma yang mengeksplorasi data, membangun model dan menemukan pola yang belum diketahui [2]. Data mining merupakan serangkaian proses untuk mengeksplorasi nilai tambah dari kumpulan data berupa pengetahuan yang tidak diketahui secara manual [3]. Data mining adalah sebuah proses dimana satu atau lebih sistem komputer pembelajarandiberi tanggung jawab untuk secara langsung memperoleh dan menganalisis pengetahuan, atau bahkan kumpulan prosedur untuk mengekstraksi nilai baru secara manual dari sekumpulan data yang tidak diketahui [4]. Data mining adalah suatu teknik yang digunakan untuk mengekstrak informasi bermanfaat dari gudang basis data yang sangat besar [5]. Data mining dapat mengatasi kumpulan data yang besar secara terbuka, sehingga tidak mungkin untuk memberikan batasan ketat pada pertanyaan yang ingin diselesaikan yang memerlukan inferensi [6].

Klasifikasi adalah proses untuk menemukan model atau fungsi yang menjelaskan dan membedakan konsep kelas data untuk memperkirakan kelas yang tidak diketahui dari suatu objek [7]. Klasifikasi merupakan teknik untuk mengidentifikasi dan mengelompokkan data berdasarkan atribut atau fitur yang dimilikinya [8]. Tujuan utama klasifikasi adalah mengidentifikasi pola dalam data dan memprediksi kelas objek yang tidak diketahui berdasarkan informasi dari objek yang diketahui [9]. Variabel kategori digunakan dalam klasifikasi target, Sebagai contoh klasifikasi keuntungan dapat dibagi menjadi tiga kategori: kecil, menengah, dan besar [10]. Dalam proses ini data mining untuk melakukan pengelompokan data terhadap bagian-bagian berdasarkan dengan atribut yang terdapat pada data [11].

Algoritma C4.5 merupakan sebuah algoritma yang dikembangkan berupa model klasifikasi dengan menggunakan teknik pohon keputusan [12]. Metode-metode atau algoritma dalam data mining memiliki banyak variasi [13]. Algoritma C4.5 yang merupakan sebuah metode pengambilan keputusan berbasis pohon keputusan sehinggadapat mengidentifikasi pola-pola kompleks dalam data [14].

Kemiskinan merupakan salah satu masalah tertinggi yang paling sering dialami oleh berbagai negara berkembang, termasuk Indonesia [15]. Bantuan sosial adalah pemberian dukungan kepada individu, keluarga, kelompok, atau komunitas dari pemerintah atau lembaga tertentu [16]. Tujuan bantuan sosial adalah untuk menuntaskan kemiskinan, yang mengandung arti bahwa bantuan sosial merupakan suatu kebijakan, program atau kegiatan masyarakat yang tidak mampu memenuhi kebutuhan dasarnya dan tidak mempunyai manfaat [17]. Bantuan Sosial biasanya diberikan oleh pemerintah atau organisasi nirlaba kepada kelompok masyarakat yang memiliki keterbatasan ekonomi atau sosial [18]. Salah satu provinsi termiskin di Indonesia adalah Sumatera Barat. Angka kemiskinan Sumatera Barat turun dari 6,55% pada tahun 2018 menjadi 5,29% pada tahun 2019 [19].

Penelitian terdahulu yang menggunakan algoritma C4.5 dalam melakukan klasifikasi penerima bantuan adalah (1) klasifikasi penerima bantuan sosial menggunakan algoritma C4.5 dalam penelitian ini memberikan nilai yang akurat dan dapat digunakan untuk pengambilan keputusan pada pihak yang berwenang. urutan atribut untuk penerima bantuan sosial ditentukan oleh pendapatan sebesar 0,227, kendaraan sebesar 0,093, jumlah tanggungan 0,0058 dan tempat tinggal 0,027. selanjutnya penelitian [20], (2) implementasi data mining menentukan penerimaan bantuan sosial pangan menggunakan algoritma C4.5 dalam penelitian ini digunakan dalam menentukan penerimaan bantuan sosial pangan di kecamatan Medan Johor dengan diterapkannya sistem ini diharapkan dapat membantu pengguna dalam menentukan penerima bantuan sosial pangan dengan

hasil yang lebih efektif dan efisien berdasarkan kriteria yang telah ditentukan [21], (3) analisis algoritma C4.5 sebagai sistem penunjang keputusan penerima bantuan kesejahteraan sosial, dalam penelitian ini menggunakan metode cross-industri standar proses. hasil analisis berdasarkan algoritma C4.5 yaitu jika usia sama dengan produktif maka layak mendapatkan bantuan [22], (4) klasifikasi kelayakan penerima bantuan sosial dengan algoritma decision tree, dalam penelitian ini menggunakan 50 dataset yang terbagi 2 jenis data diantaranya layak dan tidak layak. penelitian ini menghasilkan pohon keputusan yang dimana pendapatan lebih dari Rp.500.000 maka tidak layak mendapatkan bantuan sosial, sebaliknya jika pendapatan yang kurang dari Rp.500.000 maka layak menerima bantuan sosial [23], (5) analisis data mining menggunakan algoritma C4.5 untuk prediksi harga pasar mobil bekas. Penelitian ini menggunakan algoritma C4.5 berdasarkan data penjualan dari 2020-2022 sebanyak 50 sampel data yang menghasilkan 5 rule yang dapat menjadi acuan untuk memprediksi stabilitas harga pasar mobil bekas sehingga mendapatkan banyak keuntungan [24].

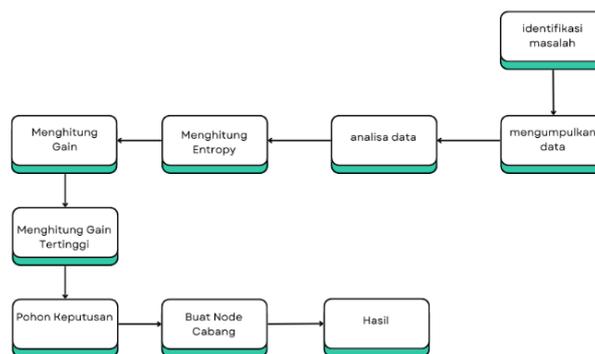
Dari penjelasan diatas dapat disimpulkan, dengan menggunakan algoritma C4.5 dalam memprediksi penerima bantuan sosial memiliki keunikan dibandingkan dengan penggunaan algoritma tersebut dalam konteks yang lain. Fokusnya adalah dapat mengurangi kemiskinan dengan cara memberikan bantuan sosial kepada yang berhak sehingga tepat sasaran. Algoritma C4.5 menganalisis penerima bantuan sosial dengan berbagai kriteria pendukung dalam menentukan penerima bantuan sosial. Penelitian ini menggunakan sampel data yang lebih banyak ketimbang penelitian sebelumnya kemudian atribut dan kriteria yang lebih kompleks dalam menentukan penerima bantuan sosial. Hasil analisis memberikan wawasan untuk pengambilan keputusan operasional dan strategis, memastikan penentuan penerima bantuan sosial lebih efisien dan lebih akurat.

## 2. Metodologi Penelitian

Dalam penelitian ini dibentuk dahulu Tahap penelitian. Tahap penelitian akan membantu dan memberikan proses apa saja dan apa yang akan dilakukan peneliti pada penelitiannya.

### 2.1. Kerangka Penelitian

Penulis melakukan tahapan penelitian adapun tahapan yang dilakukan adalah sebagai berikut:



**Gambar 1.** Kerangka Penelitian

#### a) Identifikasi Masalah

Langkah awal dari penelitian ini dimulai dari tahap Identifikasi masalah, yang diidentifikasi adalah bagaimana menerapkan algoritma C4.5 dalam melakukan prediksi Bantuan Sosial, atribut-atribut yang akan dijadikan penilaian dalam melakukan identifikasi dan bagaimana mengukur tingkat akurasi.

**b) Mengumpulkan Data**

Pengumpulan data dilakukan dengan metode studi lapangan. Penulis melakukan survey lapangan ke Kelurahan Jati Kota Padang dengan 5 atribut parameter yaitu Usia, Pekerjaan, Jumlah Anak, Kondisi Rumah Dan Status Bantuan.

**c) Analisa Data**

Data yang digunakan Data Warga Kelurahan Jati Kota Padang dengan jumlah 80 sampel. Penginputan data training ini merupakan langkah awal untuk mendapatkan nilai *Entropy* dan *Gain*.

**d) Menghitung Entropy**

Pada tahap ini melakukan proses perhitungan dengan nilai atribut sehingga menghasilkan nilai entropy. Pencarian entropy total dilakukan dengan cara mengelompokkan data dengan benar dan tepat kemudian menghitung data dengan menggunakan rumus entropy dengan persamaan (1)

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

**e) Implementasi Gain**

Dalam melakukan perhitungan *Gain*, harus mendapatkan nilai *entropy* terlebih dahulu. *Gain* adalah perbedaan *entropy* data awal dan rata-rata *entropy* setiap atribut. Dengan menggunakan rumus persamaan sebagai berikut :

$$Gain (S, A) = Entropy(s) = \sum_{i=1}^n \frac{S_i}{s} \text{Log}2 \frac{S_i}{s} \tag{2}$$

**f) Tentukan Node Tertinggi**

Pada tahap ini diperlukan untuk dapat menentukan akar dari pohon keputusan.

**g) Membuat Node Cabang**

Pada tahap ini diperlukan untuk dapat menentukan pohon keputusan

**h) Pengujian Hasil**

Proses pengujian akhir dilakukan untuk memastikan keakuratan hasil sistem dengan membandingkannya dengan data yang telah diolah secara manual. Pengujian akhir dilakukan Menggunakan Aplikasi Rapidminer Studio.

**Tabel 1** Data Bantuan Sosial

No	Nama	Usia	Pekerjaan	Jumlah Anak	Kondisi Rumah	Status Bantuan
1	Warga 1	>60	Buruh	>2	Layak	Menerima
2	Warga 2	<60	Buruh	>2	Layak	Menerima
3	Warga 3	<60	Pedagang	>2	Tidak Layak	Menerima
4	Warga 4	>60	Pedagang	>2	Tidak Layak	Menerima
5	Warga 5	>60	Pedagang	<2	Layak	Tidak Menerima
41	Warga 41	<60	Pedagang	<2	Layak	Tidak Menerima
42	Warga 42	<60	Buruh	<2	Tidak Layak	Menerima
...	...	...	..	..	..	..
76	Warga 76	>60	Belum / Tidak Bekerja	>2	Tidak Layak	Menerima
77	Warga 77	<60	Buruh	<2	Tidak Layak	Menerima
78	Warga 78	>60	Belum / Tidak Bekerja	<2	Tidak Layak	Menerima
79	Warga 79	>60	Karyawan Swasta	>2	Tidak Layak	Tidak Menerima
80	Warga 80	>60	Pedagang	>2	Layak	Menerima

## 2.2. Transformasi Data

Tabel transformasi mempunyai beberapa atribut antara lain yaitu usia, pekerjaan, penghasilan dan jumlah anak. data transformasi ini akan dijadikan dataset untuk melakukan perhitungan node. Adapun proses tersebut disajikan pada Tabel 2.

**Tabel 2** Data Transformasi

No	Usia	Pekerjaan	Jumlah anak	Kondisi Rumah	Status Bantuan
1	Lanjut	Buruh	Banyak	Layak	Menerima
2	Produktif	Buruh	Banyak	Layak	Menerima
3	Produktif	Pedagang	Banyak	Tidak Layak	Menerima
4	Lanjut	Pedagang	Banyak	Tidak Layak	Menerima
5	Lanjut	Pedagang	Sedikit	Layak	Tidak Menerima
49	Produktif	Pedagang	Sedikit	Layak	Tidak menerima
50	Produktif	Buruh	Sedikit	Tidak Layak	Menerima
...	.....	.....	.....	.....	.....
76	Lanjut	Belum / Tidak Bekerja	Banyak	Layak	Menerima
77	Produktif	Buruh	Sedikit	Tidak Layak	Menerima
78	Lanjut	Belum / Tidak Bekerja	Sedikit	Tidak Layak	Menerima
79	Lanjut	Karyawan Swasta	Banyak	Tidak Layak	Tidak Menerima
80	Lanjut	Petani / Berkebun	Banyak	Layak	Menerima

## 3. Hasil dan Pembahasan

### 3.1. Perhitungan Entropy Total

Pencarian *entropy* total dan *gain* dilakukan dengan cara mengelompokkan data dengan benar dan tepat. Kemudian menghitung data menggunakan rumus *entropy* (1) dan rumus *Gain* (2) pada masing-masing atribut data.

$$Entropy\ total = \left( \left( \frac{19}{80} * \left( \log_2 \left( \frac{19}{80} \right) \right) \right) + \left( \left( -\frac{61}{80} * \left( \log_2 \left( \frac{61}{80} \right) \right) \right) \right) = 0,790858091$$

### 3.2. Menghitung Entropy Setiap Atribut

Perhitungan nilai *entropy* pada setiap atribut (usia, Pekerjaan, Jumlah Anak dan Kondisi Rumah) dengan menggunakan persamaan (1) dapat dilihat sebagai berikut :

#### a) Atribut Usia

*Entropy* kategori produktif

$$Entropy = \left( \left( -\frac{10}{46} * \left( \log_2 \left( \frac{10}{46} \right) \right) \right) + \left( \left( -\frac{36}{46} * \left( \log_2 \left( \frac{36}{46} \right) \right) \right) \right) = 0,755375413$$

*Entropy* Kategori Lanjut

$$Entropy = \left( \left( -\frac{9}{34} * \left( \log_2 \left( \frac{9}{34} \right) \right) \right) + \left( \left( -\frac{25}{34} * \left( \log_2 \left( \frac{25}{34} \right) \right) \right) \right) = 0,833764907$$

#### b) Atribut Pekerjaan

*Entropy* kategori buruh

$$Entropy = \left( \left( -\frac{2}{19} * \left( \log_2 \left( \frac{2}{19} \right) \right) \right) + \left( \left( -\frac{17}{19} * \left( \log_2 \left( \frac{17}{19} \right) \right) \right) \right) = 0,485460761$$

*Entropy* kategori petani / berkebun

$$Entropy = \left( \left( -\frac{1}{15} * \left( \log_2 \left( \frac{1}{15} \right) \right) \right) + \left( \left( -\frac{14}{15} * \left( \log_2 \left( \frac{14}{15} \right) \right) \right) \right) = 0,353359335.$$

*Entropy* kategori Karyawan Swasta

$$Entropy = \left( \left( -\frac{12}{12} * \left( \log_2 \left( \frac{12}{12} \right) \right) \right) + \left( \left( -\frac{0}{12} * \left( \log_2 \left( \frac{0}{12} \right) \right) \right) \right) = 0$$

Entropy Kategori Pedagang

$$Entropy = \left( \left( -\frac{4}{19} * \left( \log_2 \left( \frac{4}{19} \right) \right) \right) + \left( \left( -\frac{15}{19} * \left( \log_2 \left( \frac{15}{19} \right) \right) \right) \right) \right) = 0,74248757.$$

Entropy kategori tidak bekerja

$$Entropy = \left( \left( -\frac{15}{15} * \left( \log_2 \left( \frac{15}{15} \right) \right) \right) + \left( \left( -\frac{0}{15} * \left( \log_2 \left( \frac{0}{15} \right) \right) \right) \right) \right) = 0$$

**c) Atribut Jumlah Anak**

Entropy kategori Sedikit.

$$Entropy = \left( \left( -\frac{8}{36} * \left( \log_2 \left( \frac{8}{36} \right) \right) \right) + \left( \left( -\frac{28}{36} * \left( \log_2 \left( \frac{28}{36} \right) \right) \right) \right) \right) = 0,764204507$$

Entropy kategori Banyak.

$$Entropy = \left( \left( -\frac{11}{44} * \left( \log_2 \left( \frac{11}{44} \right) \right) \right) + \left( \left( -\frac{33}{44} * \left( \log_2 \left( \frac{33}{44} \right) \right) \right) \right) \right) = 0,811278124$$

**d) Atribut Kondisi Rumah**

Entropy kategori Layak.

$$Entropy = \left( \left( -\frac{9}{34} * \left( \log_2 \left( \frac{9}{34} \right) \right) \right) + \left( \left( -\frac{25}{34} * \left( \log_2 \left( \frac{25}{34} \right) \right) \right) \right) \right) = 0,833764907$$

Entropy kategori Tidak Layak.

$$Entropy = \left( \left( -\frac{10}{46} * \left( \log_2 \left( \frac{10}{46} \right) \right) \right) + \left( \left( -\frac{36}{46} * \left( \log_2 \left( \frac{36}{46} \right) \right) \right) \right) \right) = 0,7553375413$$

### 3.3. Menghitung Gain

Perhitungan nilai *Gain* dilakukan pada setiap atribut yaitu (usia, pekerjaan, jumlah anak dan Kondisi Rumah) informasi lebih lanjut mengenai perhitungan nilai *gain* dapat dilihat pada persamaan (2). Hasil perhitungan nilai *gain* untuk masing-masing atribut dapat ditentukan dalam konteks berikut ini:

**a) Atribut Usia**

$$Gain\ Usia = \left( 0,790858091 - \left( \frac{46}{80} * 0,7553375413 \right) + \left( \frac{34}{80} * 0,833764907 \right) \right) = 0,002167143.$$

**b) Atribut Pekerjaan**

$$Gain\ Pekerjaan = \left( 0,790858091 - \left( \frac{19}{80} * 0,485460761 \right) + \left( \frac{15}{80} * 0,353359335 \right) + \left( \frac{12}{80} * 0 \right) + \left( \frac{19}{80} * 0,74248757 \right) + \left( \frac{15}{80} * 0 \right) \right) = 0,432965487$$

**c) Atribut Jumlah Anak**

$$Gain\ Jumlah\ Anak = \left( 0,790858091 - \left( \frac{36}{80} * 0,764204507 \right) + \left( \frac{44}{80} * 0,811278124 \right) \right) = 0,355886141$$

**d) Atribut Kondisi Rumah**

$$Gain\ Kondisi\ Rumah = \left( 0,790858091 - \left( \frac{34}{80} * 0,833764907 \right) + \left( \frac{46}{80} * 0,7553375413 \right) \right) = 0,002167143$$

### 3.4. Perhitungan Node

Nilai *Entropy* dan *gain* dapat diperoleh dengan menggunakan persamaan (1) dan (2). Informasi lebih tentang nilai-nilai tersebut dapat ditemukan pada Tabel 3 dibawah ini. Tabel 3 berisi data mengenai hasil perhitungan *entropy* dan *gain*.

**Tabel 3.** Perhitungan Node 1

Node 1	Jumlah (S)	Tidak Menerima	Menerima	Entropy	Gain
Total	80	19	61	0,790858	
Usia					0,0021

Node 1	Jumlah (S)	Tidak Menerima	Menerima	Entropy	Gain
Produktif	46	10	36	0,7553754	
Lanjut	34	9	25	0,8337649	
<b>Pekerjaan</b>					0,43296
Buruh	19	2	17	0,485460	
Petani / berkebun	15	1	14	0,3533593	
Karyawan Swasta	12	12	0	0	
Pedagang	19	4	15	0,74248757	
Tidak bekerja	15	0	15	0	
<b>Jumlah Anak</b>					0,35588
Sedikit	36	9	28	0,7642045	
Banyak	44	11	33	0,8112781	
<b>Kondisi Rumah</b>					0,00216
Layak	34	9	25	0,8337649	
Tidak Layak	46	10	36	0,7553754	

### 3.4.1. Perhitungan Node 1.1

Perhitungan *node* 1.1 dengan menggunakan persamaan 1 dan 2, dilakukan untuk menentukan *node* selanjutnya, dapat dilihat pada Tabel 4 berikut ini.

Tabel 4. Perhitungan Node 1.1

Node 1	Jumlah (S)	Tidak Menerima	Menerima	Entropy	Gain
Total	19	4	15	0,7424875	
<b>Usia</b>					0,08617
Produktif	11	1	10	0,4394969	
Lanjut	8	3	5	0,9544340	
<b>Jumlah Anak</b>					0,02420
Sedikit	11	3	8	0,8453509	
Banyak	8	1	7	0,5435644	
<b>Kondisi Rumah</b>					0,02420
Layak	11	3	8	0,8453509	
Tidak Layak	8	1	7	0,5435644	

### 3.4.2. Perhitungan Node 1.1.1

Perhitungan *node* 1.1.1 dengan menggunakan persamaan 1 dan 2 dilakukan untuk menentukan *node* selanjutnya, dapat dilihat pada Tabel 5 berikut ini.

Tabel 5. Perhitungan Node 1.1.1

Node 1	Jumlah (S)	Tidak Menerima	Menerima	Entropy	Gain
Total	34	9	25	0,8337649	
<b>Jumlah Anak</b>					0,00321
Sedikit	17	5	12	0,8739810	
Banyak	14	4	13	0,7871265	
<b>Kondisi Rumah</b>					0,00113
Layak	14	4	10	0,8631205	
Tidak Layak	20	5	15	0,8112781	

### 3.4.3. Perhitungan Node 1.1.1.1

Perhitungan *node* 1.1.1.1 dengan menggunakan persamaan 1 dan 2 dilakukan untuk menentukan *node* selanjutnya, dapat dilihat pada Tabel 6 berikut ini.

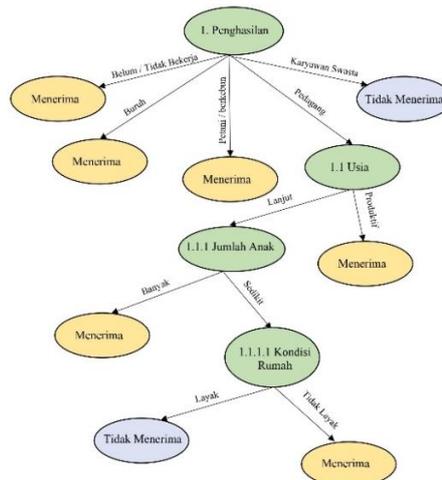
Tabel 6. Menghitung Node 1.1.1.1

Node 1	Jumlah (S)	Tidak Menerima	Menerima	Entropy	Gain
Total	36	8	28	0,7642045	

Node 1	Jumlah (\$)	Tidak Menerima	Menerima	Entropy	Gain
<b>Kondisi Rumah</b>					0,13899
Layak	9	5	4	0,991076	
Tidak Layak	27	3	24	0,5032583	

### 3.5. Hasil Pohon Keputusan

Pohon keputusan merupakan hasil akhir dari kinerja metode C4.5. hasil tersebut digambarkan berdasarkan pembahasan sebelumnya. Adapun hasil keluaran pohon keputusan dapat disajikan pada Gambar 2.

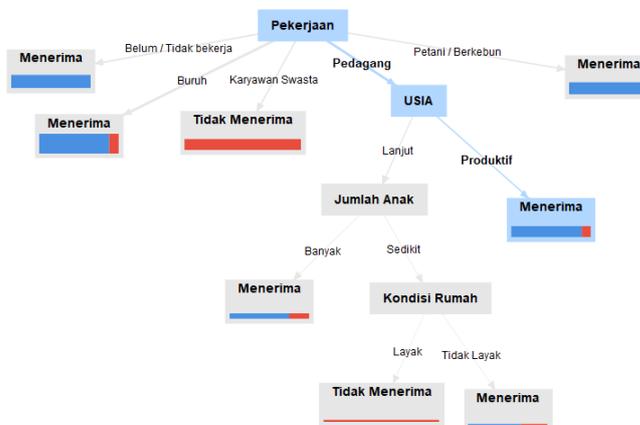


Gambar 2. Pohon Keputusan

Dari Gambar 2 menghasilkan aturan jika Pekerjaan pedagang dan usia lanjut kemudian jumlah anak sedikit dan kondisi rumah layak maka status bantuan tidak menerima tetapi jika pekerjaan pedagang dan usia lanjut kemudian jumlah anak sedikit dan kondisi rumah tidak layak maka status bantuan menerima bantuan sosial.

### 3.6. Hasil Pohon Keputusan Menggunakan Rapidminer

Setelah melakukan pengimputan data penerima bantuan sosial ke aplikasi *rapidminer* dengan mengikuti langkah-langkah yang sudah ditentukan maka didapatkan hasil pohon keputusan seperti Gambar 2.



Gambar 3. Pohon Keputusan Menggunakan Rapidminer

Gambar 3 merupakan hasil kinerja metode yang memberikan Gambaran dalam pohon keputusan. Berdasarkan pohon keputusan yang telah dibentuk, maka dapat dibuatkan *rule* atau aturan. Aturan tersebut dapat disajikan diantaranya :

- 1) *IF* Pekerjaan = Buruh *THEN* Status Bantuan Menerima
- 2) *IF* Pekerjaan = Karyawan Swasta *THEN* Status Bantuan Tidak Menerima.
- 3) *IF* Pekerjaan = Petani / Berkebun *THEN* Status Bantuan Menerima.
- 4) *IF* Pekerjaan = *Belum / Tidak Bekerja* *THEN* Status Bantuan Menerima.
- 5) *IF* Pekerjaan = Pedagang *AND* Usia = Produktif *THEN* Status Bantuan Menerima.
- 6) *IF* Pekerjaan = Pedagang *AND* Usia = Lanjut *AND* Jumlah anak = Banyak *THEN* Status Bantuan Menerima.
- 7) *IF* Pekerjaan = Pedagang *AND* Usia = Lanjut *AND* Jumlah anak = Sedikit *AND* Kondisi Rumah = Layak *THEN* Status Bantuan Tidak Menerima.
- 8) *IF* Pekerjaan = Pedagang *AND* Usia = Lanjut *AND* Jumlah anak = Sedikit *AND* Kondisi Rumah = Tidak Layak *THEN* Status Bantuan Menerima.

Berdasarkan *rule* yang telah terbentuk sebanyak 8 *rule*, *rule* menggambarkan suatu pengetahuan yang telah diperoleh dari kumpulan data penerima bantuan sosial.

#### 4. Kesimpulan

Berdasarkan analisis data mining menggunakan algoritma c4.5 dalam memprediksi penerima bantuan sosial di kelurahan jati kota padang menghasilkan akurasi yaitu sebesar 93.75%. Hasil tersebut didapatkan berdasarkan data penerima bantuan sosial selama Selama tahun 2024 sebanyak 80 data sampel yang menghasilkan 8 *rule*. Proses kinerja metode dilakukan dengan pengujian menggunakan *rapidminer* menghasilkan aturan dalam penentuan penerima bantuan sosial. Hasil tersebut dapat digunakan untuk menjadikan acuan dalam menentukan penerima bantuan sosial selanjutnya yang lebih akurat dan lebih adil agar masyarakat yang lebih berhak mendapatkan bantuan sosial.

#### Daftar Pustaka

- [1] Boy Sandy Dwi Nugraha.H, S. Defit, And G. Widi Nurcahyo, “Algoritma C4.5 Untuk Prediksi Bimbingan Siswa Berdasarkan Tipologi Hippocrates-Galenus,” Jurnal Teknoif Teknik Informatika Institut Teknologi Padang, Vol. 11, No. 1, Pp. 1–8, 2023, Doi: 10.21063/Jtif.2023.V11.1.1-8.
- [2] F. Pirmansyah And T. Wahyudi, “Implementasi Data Mining Menggunakan Algoritma C4.5 Untuk Prediksi Evaluasi Anggota Satuan Pengamanan Studi Kasus Pt. Yimm Pulogadung,” Jurnal Indonesia : Manajemen Informatika Dan Komunikasi, Vol. 4, No. 3, Pp. 1566–1580, Sep. 2023, Doi: 10.35870/Jimik.V4i3.370.
- [3] A. Nadira And F. Utami, “Implementasi Data Mining Dalam Mengidentifikasi Faktor Pasien Yang Berpotensi Mengalami Obesitas Menggunakan Algoritma C4.5,” 2024.
- [4] M. Roghib, N. Rahaningsih, And R. D. Dana, “Penerapan Algoritma C4.5 Untuk Seleksi Penjurusan Siswa Baru Pada Sekolah Menengah Kejuruan (Studi Kasus: Smk Plus Al-Hilal Arjawinangun),” 2024.
- [5] D. Ruswanti, D. Susilo, And R. Riani, “Implementasi Crisp-Dm Pada Data Mining Untuk Melakukan Prediksi Pendapatan Dengan Algoritma C.45,” Go Infotech: Jurnal Ilmiah Stmik Aub, Vol. 30, No. 1, Pp. 111–121, Jun. 2024, Doi: 10.36309/Goi.V30i1.266.
- [6] M. Anjas Aprihartha Et Al., “Penyelesaian Masalah Ketidakseimbangan Data Melalui Teknik Oversampling Dan Undersampling Pada Klasifikasi Siswa Tidak

- Naik Kelas,” *Jurnal Teknik Ibnu Sina*, Vol. 9, No. 01, 2024, Doi: 10.36352/Jt-Ibsi.V9i01.807.
- [7] Y. Mulyanto, F. Idifitriani, And A. Wati, “Klasifikasi Data Mining Untuk Penentuan Stunting Pada Balita Menggunakan Metode Naïve Bayes,” 2024.
- [8] A. R. Pradih Et Al., “Perbandingan Algoritma Klasifikasi Data Mining Untuk Prediksi Kualitas Udara Di Kota Bandung,” 2024.
- [9] D. M. Musa Et Al., “Penerapan Data Mining Untuk Klasifikasi Data Penjualan Pakan Ternak Terlaris Dengan Algoritma C4.5,” *Jurnal Teknologi Informatika Dan Komputer*, Vol. 10, No. 1, Pp. 168–182, Mar. 2024, Doi: 10.37012/Jtik.V10i1.1985.
- [10] A. Julianto And S. Andayani, “Penerapan Data Mining Untuk Klasifikasi Produk Terlaris Menggunakan Algoritma Naive Bayes Pada Bengkel Motor,” *Jusitik : Jurnal Sistem Dan Teknologi Informasi Komunikasi*, Vol. 7, No. 2, Pp. 50–58, Jul. 2024, Doi: 10.32524/Jusitik.V7i2.1148.
- [11] D. Librado And A. H. Nasyuha, “Penerapan Data Mining Untuk Klasifikasi Penerima Kredit Dengan Perbandingan Algoritma Naïve Bayes Dan Algoritma C4.5,” *Jurnal Media Informatika Budidarma*, Vol. 7, No. 4, P. 1952, Oct. 2023, Doi: 10.30865/Mib.V7i4.6907.
- [12] R. Musfekar, H. Apriadinata, And B. Yusuf, “Aplikasi Prediksi Prestasi Pada Siswa Menggunakan Algoritma C4.5,” *Jurnal Manajemen Informatika (Jamika)*, Vol. 13, No. 2, Pp. 148–162, Oct. 2023, Doi: 10.34010/Jamika.V13i2.10649.
- [13] Dyah Ardyani Rizqi Azizah Adha, Aulia Noveesa Allanda, Diah Ayu Fatmasari, And Siska Narulita, “Performansi Algoritma C4.5 Untuk Prediksi Kelulusan Mahasiswa,” *Jurnal Cakrawala Informasi*, Vol. 3, No. 2, Pp. 9–17, Dec. 2023, Doi: 10.54066/Jci.V3i2.339.
- [14] C. A. Cahyani Et Al., “Penerapan Algoritma C4.5 Terhadap Klasifikasi Kepuasan Penumpang Pada Maskapai Penerbangan,” 2024. [Online]. Available: [www.kaggle.com](http://www.kaggle.com)
- [15] W. W. Arupandani, F. Taufik, And R. Mahyuni, “Implementasi Data Mining Menentukan Penerimaan Bantuan Sosial Pangan (Bsp) Menggunakan Algoritma C4.5,” *Jurnal Sistem Informasi Triguna Dharma (Jursi Tgd)*, Vol. 2, No. 5, P. 705, 2023, Doi: 10.53513/Jursi.V2i5.5612.
- [16] S. Suprpto, E. Edora, And F. A. Pasaribu, “Sistem Pendukung Keputusan Calon Penerima Program Bantuan Sosial (Bansos) Menggunakan Metode Simple Additive Weighting (Saw),” *Malcom: Indonesian Journal Of Machine Learning And Computer Science*, Vol. 4, No. 1, Pp. 188–197, 2024, Doi: 10.57152/Malcom.V4i1.1057.
- [17] C. Dian, V. Zalukhu, F. Izwar, N. Aprilian, And R. Dwi, “Pengaruh Program Bansos Terhadap Tingkat Kemiskinan Kota Padang,” Vol. 3, No. 1, Pp. 87–92, 2024.
- [18] A. P. Pristiawati, I. Permana, Z. Zarnelly, And F. Muttakin, “Klasifikasi Penerima Bantuan Beras Miskin Menggunakan Algoritma K-Nn, Nbc Dan C4.5,” *Building Of Informatics, Technology And Science (Bits)*, Vol. 5, No. 1, Jun. 2023, Doi: 10.47065/Bits.V5i1.3617.
- [19] C. Dian Veronika Zalukhu, F. Izwar, N. Aprilian, R. Dwi Cahyani, And S. Mishwa Hanifah, “Pengaruh Program Bansos Terhadap Tingkat Kemiskinan Kota Padang,” Vol. 3, No. 1, Doi: 10.38035/Jim.V3i1.
- [20] A. Junaidi, S. Agustiani, C. Ika Agustyaningrum, Y. Tajul Arifin, And U. Bina Sarana Informatika, “This Work Is Licensed Under A Creative Commons Attribution-Sharealike 4.0 International License Jurnal Teknik Komputer Amik Bsi Klasifikasi Penerima Bantuan Sosial Menggunakan Algoritma C 4.5”, Doi: 10.31294/Jtk.V4i2.

- [21] W. W. Arupandani, F. Taufik, And R. Mahyuni, "Implementasi Data Mining Menentukan Penerimaan Bantuan Sosial Pangan (Bsp) Menggunakan Algoritma C4.5", [Online]. Available: <https://ojs.trigunadharma.ac.id/index.php/jsi>
- [22] B. Basiroh And H. Irfananto, "C4.5 Algorithm As A Decision Support System For Social Welfare Aid Recipients," *Jurnal Teknologi Informasi Dan Komunikasi (Tikomsin)*, Vol. 12, No. 1, P. 9, Apr. 2024, Doi: 10.30646/Tikomsin.V12i1.816.
- [23] F. Rahmadayanti, A. Lovita, And S. Muntari, "Algoritma Decision Tree".
- [24] R. N. Fazira, A. Wanto, I. Gunawan, S. Tunas Bangsa, And G. Artikel, "Pemanfaatan Algoritma C4.5 Untuk Klasifikasi Siswa Yang Layak Di Ajukan Dalam Memperoleh Pip Utilization Of The C4.5 Algorithm For Classifying Students Who Are Eligible To Apply For Obtaining Pip," *Jomlai: Journal Of Machine Learning And Artificial Intelligence*, Vol. 2, No. 2, Pp. 2828–9099, 2023, Doi: 10.55123/Jomlai.V2i2.3187.