
ALGORITMA KLASIFIKASI DATA MINING UNTUK MEMREDIKSI MASYARAKAT DALAM MENERIMA BANTUAN SOSIAL

Ardi Ramdani^a, Christian Dwi Sofyan^b, Fauzi Ramdani^c, Muhamad Fauzi Arya Tama^d, Muhammad Angga Rachmatsyah^e

^{a-e}Teknik Informatika, Universitas Adhirajasa Reswara Sanjaya

Email: anggarachmatsyah@gmail.com

ARTICLE INFO

Article history:

Received : xx – xxxx - 2021

Received in revised form : xx – xxxx - 2021

Accepted : xx – xxxx - 2021

Available online : xx – xxxx - 2021

ABSTRACT

Social assistance is one of the government programs for the welfare of the community in order to make a better life. Especially for people who are economically disadvantaged. However, not everyone will get social assistance due to several factors. To overcome the existing problems, the government held a social assistance program for the underprivileged in economic terms. In this study, the author uses the Datamining Classification Algorithm, namely the C4.5 Algorithm to predict the community in obtaining social funding assistance. The data sample was taken from the data of recipients of social assistance in the Binong sub-district, Subang district. From the test results, the Cross Validation and Confusion Matrix tests. The results obtained for the Accuracy Algorithm C4.5 is 100.00%, the value for Precision is 100.00%, and the value for Sensitivity or Recall is 100.00%. Thus the C4.5 Algorithm is the best algorithm and technique for Predicting the Community in obtaining Social Fund Assistance.

Keywords: data mining, classification algorithm, C4.5 algorithm, Social Assistance..

ABSTRAK

Bantuan sosial merupakan salah satu program pemerintah untuk mensejahterakan masyarakat agar menjadikan kehidupan yang lebih baik. Terutama bagi masyarakat yang kurang mampu dalam hal ekonomi. Akan tetapi tidak semua orang akan mendapatkan bantuan sosial karena beberapa faktor penyebab. Untuk mengatasi permasalahan yang ada maka pemerintah mengadakan program bantuan sosial bagi masyarakat kurang mampu dalam hal ekonomi. Pada penelitian ini penulis menggunakan Algoritma Klasifikasi Datamining yaitu Algoritma C4.5 untuk memprediksi masyarakat dalam memperoleh bantuan dana sosial. Sampel data diambil dari data penerima bantuan sosial kecamatan binong, kabupaten Subang. Dari hasil pengujian digunakan tes Cross Validation dan Confusion Matrix. Hasil yang diperoleh untuk nilai Accuracy Algoritma C4.5 adalah sebesar 100.00%, nilai untuk Precision sebesar 100.00%, dan nilai untuk Sensitivity atau Recall sebesar 100.00%. Dengan demikian Algoritma C4.5 merupakan algoritma dan teknik terbaik untuk Memprediksi Masyarakat dalam memperoleh Bantuan Dana Sosial.

Kata Kunci : data mining, algoritma klasifikasi, algoritma C4.5, Bantuan Sosial.

Pendahuluan

Bantuan sosial adalah bantuan yang diberikan oleh pemerintah daerah kepada masyarakat dalam bentuk dana atau kebutuhan hidup. Sifat bantuan ini adalah untuk jangka waktu tertentu, tidak terus menerus [1]. Bantuan ini dipilih secara selektif dalam bentuk mata uang atau barang, dan penyalurannya disesuaikan dengan kemampuan keuangan daerah. Tujuannya adalah untuk mendukung program dan kegiatan pemerintah daerah untuk membuat kehidupan masyarakat lebih baik dan lebih sejahtera dengan mengutamakan prinsip keadilan, kebersamaan, dan kepentingan masyarakat. Jaminan sosial ini diatur dalam Undang-Undang Nomor 40 Tahun 2004 tentang Sistem Jaminan Sosial Nasional. Garis besarnya merupakan sebuah bidang kesejahteraan sosial yang di dalamnya memperhatikan perlindungan sosial termasuk kemiskinan, usia lanjut, kecacatan, pengangguran, keluarga dan anak-anak. Bantuan sosial sangat berpengaruh terhadap laju pertumbuhan penduduk di Indonesia [2]. Namun sudah tentu syarat dan ketentuan berlaku, dan semua kriteria pemilihan penentuan masyarakat dalam memperoleh bantuan sosial itu dilakukan berdasarkan data yang ada, kemudian dianalisis, namun terkadang hasil yang diperoleh tidak sesuai. Dengan demikian pada penelitian ini penulis menggunakan Algoritma Klasifikasi Data mining yaitu Algoritma C4.5 untuk memprediksi masyarakat dalam memperoleh bantuan dana pendidikan. Sampel data diambil dari data penerima bantuan sosial kecamatan binong, kabupaten Subang.

Algoritma C4.5 merupakan algoritma yang digunakan untuk membentuk pohon keputusan. *Decision Tree* adalah metode yang sangat kuat dan terkenal untuk klasifikasi dan prediksi. Metode *Decision Tree* mengubah fakta yang sangat besar menjadi *Decision Tree* yang mewakili aturan. Aturan dapat dengan mudah dipahami dalam bahasa alami. Dan juga dapat direpresentasikan dalam bentuk bahasa database seperti *Structured Query Language* untuk mencari *record* dalam kategori tertentu. *Decision Tree* juga berguna untuk mengeksplorasi data, menemukan hubungan tersembunyi antara beberapa variabel input kandidat dan variabel target [3].

Penelitian ini dilakukan dengan menggunakan Algoritma C4.5 dengan menganalisis atribut-atribut untuk dijadikan parameter untuk memprediksi masyarakat dalam penerimaan bantuan sosial. Parameter itu terdiri dari nama, alamat, penghasilan, pekerjaan, kepemilikan rumah, status dan kelayakan dalam menerima bantuan sosial.

TINJAUAN PUSTAKA

Data Mining

Data mining telah menjadi metode yang sangat berguna karena kemampuannya untuk mengubah sejumlah besar data menjadi informasi dan pengetahuan yang berguna. Informasi dan pengetahuan yang diperoleh dapat digunakan untuk menganalisis berbagai objek, seperti analisis pasar, deteksi penipuan dan retensi pelanggan, untuk pengendalian produksi dan ilmu eksplorasi [4].

Definisi lain dari *Data mining* adalah proses mengekstraksi data yang berguna dari. Basis pengetahuan yang dikombinasikan dengan keterampilan dan keahlian analitis tingkat lanjut dapat memainkan peran penting dalam mengungkap pola dan tren tersembunyi. Data ini dapat diabstraksikan menggunakan banyak algoritma data mining yang berbeda. *Data Mining* memainkan peran yang sangat penting dalam mengintegrasikan teknologi dan penelitian. Identifikasi aturan asosiasi dengan klasifikasi. Pengenalan bekerja dengan kategorisasi, yang menangani beberapa hasil buruk, rata-rata, dan baik[5].

Data Klasifikasi

Klasifikasi adalah teknik klasifikasi adalah salah satu teknik data mining yang paling banyak diterapkan dalam berbagai aplikasi. Proses klasifikasi membutuhkan dua jenis data: data pelatihan dan data pengujian. Data pelatihan adalah data yang digunakan oleh algoritma data mining untuk mempelajari metrik klasifikasi untuk mengklasifikasikan data lain yaitu data pengujian. Dua set data artikel teks digunakan dan diklasifikasikan menjadi data pelatihan dan data pengujian[6].

Algoritma C4.5

Algoritma C4.5 merupakan algoritma yang digunakan untuk membentuk pohon keputusan [3]. Algoritma C4.5 merupakan teknik *data mining* yang dapat membuat prediksi dengan mengolah variabel data [7]. Tujuan dari algoritma C4.5 ini adalah untuk mengeksplorasi data dan menemukan beberapa hubungan tersembunyi antara beberapa variabel input kandidat dan variabel target [8]. Banyak algoritma yang tersedia untuk pembentukan pohon keputusan, termasuk: ID3, CART (*Classification and Regression Tree*), C4.5, dan C5.0 [9]. Algoritma C4.5 merupakan pengembangan dari algoritma ID3[10]. Proses dari algoritma C4.5 adalah mengubah bentuk data (tabel) menjadi model pohon, mengubah model pohon menjadi aturan, dan menyederhanakan aturan[8].

Tahapan algoritma C4.5[3] adalah sebagai berikut:

1. Siapkan data pelatihan (*training*).
2. Pilih atribut sebagai akar

Untuk memilih atribut akar, dilihat pada nilai Gain tertinggi dari atribut-atribut yang ada. Cara mendapatkan Nilai Gain yaitu dengan mencarinya dengan rumus Nilai. Rumus Entropi :

❖ Untuk perhitungan nilai Entropy sbb :

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

▪ Keterangan :

- S : himpunan kasus.
- A : fitur.
- n : jumlah partisi S.
- pi : proporsi dari S_i terhadap S

Untuk mendapatkan nilai Gain bisa menggunakan rumus sebagai berikut :

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

Decision Tree

Decision tree merupakan metode pembelajaran terawasi non-parametrik yang digunakan untuk klasifikasi dan regresi. Tujuannya adalah untuk membuat model yang memprediksi nilai variabel target dengan mempelajari aturan keputusan sederhana yang diturunkan dari fitur data. Sebuah pohon dapat dilihat sebagai pendekatan konstan sepotong demi sepotong. Ini adalah diagram seperti flowchart yang menunjukkan berbagai hasil dari serangkaian keputusan. Ini dapat digunakan sebagai alat pengambilan keputusan, untuk analisis penelitian, atau untuk strategi perencanaan. Keuntungan utama dari pohon keputusan adalah dapat dicocokkan dengan kumpulan data dengan cepat dan model akhir dapat divisualisasikan dan diinterpretasikan dengan rapi menggunakan diagram “pohon”. Kerugian utama adalah bahwa pohon keputusan cenderung terlalu sesuai dengan kumpulan data pelatihan, yang berarti kemungkinan berkinerja buruk pada data yang tidak terlihat. Itu juga dapat sangat dipengaruhi oleh outlier dalam dataset [11].

Cross Validation

Cross Validation merupakan salah satu teknik untuk mengevaluasi nilai keakuratan model yang dirancang berdasarkan dataset tertentu. Data yang digunakan dalam proses pembangunan model disebut data training, sedangkan data yang akan digunakan untuk memvalidasi model disebut sebagai data testing [12].

Tabel 1. *Confusion Matrix*

Kelas	Prediksi Yes	Prediksi No	Total
Aktual Yes	True Positive (TP)	False Negative (FN)	Positive (P)
Aktual No	False Positive (FP)	True Negative (TN)	Negative (N)
Total	P'	N'	P+N

Rumus Mencari Akurasi adalah :

$$accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

Rumus Mencari *Sensitivity* atau *Recall* adalah :

$$sensitivity = \frac{TP}{TP + FN}$$

Rumus Mencari Nilai *Precision* adalah :

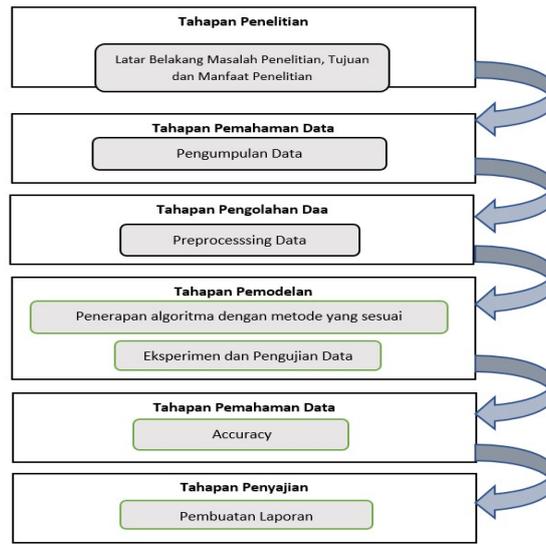
$$precision = \frac{TP}{TP + FP}$$

Rapidminer

Rapidminer merupakan *software* yang bersifat *open source* (sumber terbuka). RapidMiner adalah solusi untuk penelitian data mining, penelitian text mining, dan penelitian prediktif. RapidMiner menggunakan berbagai teknik deskriptif dan prediktif untuk memberikan wawasan kepada user sehingga mereka dapat membuat keputusan terbaik. RapidMiner memiliki sekitar 500 operator data mining, termasuk operator untuk masukan, keluaran, *preprocessing* data, dan visualisasi. RapidMiner adalah perangkat lunak analisis data yang berdiri sendiri yang bertindak sebagai mesin *data mining* dan dapat diintegrasikan ke dalam produknya sendiri. RapidMiner ditulis dalam bahasa Java, sehingga berjalan di semua sistem operasi [12].

METODE PENELITIAN

Penelitian yang digunakan adalah model penelitian eksperimen. Bertujuan untuk melakukan prediksi Masyarakat dalam memperoleh Bantuan Dana dari pemerintah, berdasarkan nilai akurasi dan evaluasi pada algoritma klasifikasi data mining. Penelitian ini menekankan pada teori-teori yang sudah ada dengan dilandasi oleh kerangka pemikiran pemecahan masalah seperti pada gambar 1 dibawah.



Gambar 1. Kerangka Penelitian

Penelitian ini dilakukan dengan menjalankan beberapa langkah proses penelitian yaitu :

1. Pengumpulan Data.
2. Pengolahan Data Awal.
3. Pengukuran Penelitian.
4. Analisa hasil Penerapan Algoritma.

HASIL DAN PEMBAHASAN

Pengumpulan Data

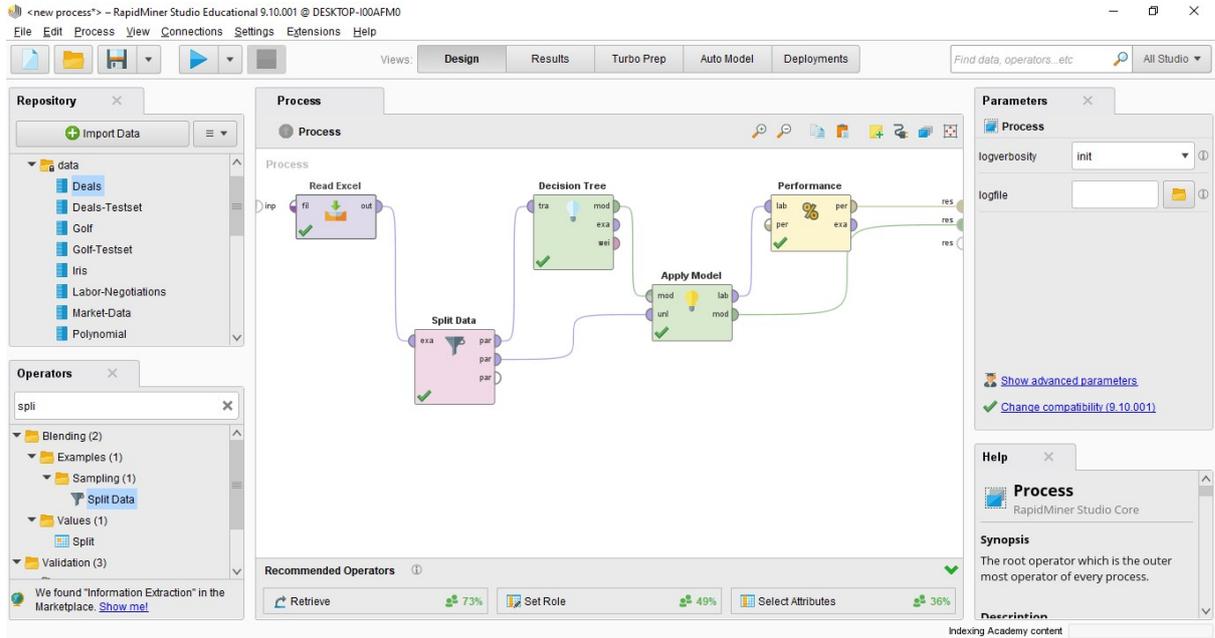
Data yang digunakan bersumber dari data Masyarakat Kecamatan Binong, Kabupaten Subang seperti terlihat dalam gambar 2 dibawah ini:

No	Nama	Alamat	Penghasilan	Pekerjaan	Kepemilikan Rumah	Status	Kelayakan
1	AMINAH	DUSUN CIBARENGKOK RT 009 RW 003	lebih dari UMKM	wiraswasta	Milik Pribadi	aktif	tidak layak
2	KARNESIH	CIGADING RT 018 RW 011	lebih dari UMKM	wiraswasta	Milik Pribadi	aktif	tidak layak
3	TRIAH JULIA	KAMPUNG BABAKAN RESIK RT 001 RW 001	lebih dari UMKM	wiraswasta	Milik Pribadi	aktif	tidak layak
4	NENCE	DUSUN KRAJAN II RT 006 RW 002	kurang dari UMKM	karyawan	Kontrak	aktif	layak
5	NURAEINI	MUNGKUR JAYA RT 012 RW 003	kurang dari UMKM	karyawan	kontrak	aktif	layak
6	ELA KOMALASARI	DUSUN WATES RT 015 RW 004	kurang dari UMKM	karyawan	Kontrak	aktif	layak
7	HENI SUHAENI	KAMPUNG MUNGKUR JAYA RT 012 RW 003	lebih dari UMKM	wiraswasta	Milik Pribadi	aktif	tidak layak
8	ARI APRILIAYANI	DUSUN KRAJAN II RT 007 RW 002	lebih dari UMKM	wiraswasta	Milik Pribadi	aktif	tidak layak
9	ABDUL OMAN	KAMPUNG BABAKAN RESIK RT 001 RW 001	kurang dari UMKM	karyawan	Milik Pribadi	aktif	layak
10	IIS TETI	KAMPUNG BABAKAN RESIK RT 001 RW 001	kurang dari UMKM	karyawan	Milik Pribadi	aktif	layak
11	ELU LUARISH	DUSUN BABAKAN SAPI RT 008 RW 002	kurang dari UMKM	wiraswasta	Milik Pribadi	aktif	layak
12	ERAH HERAWATI	KAMPUNG BABAKAN SAPI RT 008 RW 002	kurang dari UMKM	karyawan	Milik Pribadi	aktif	layak
13	SARIYAH	DUSUN MUNGKUR JAYA RT 012 RW 003	kurang dari UMKM	karyawan	Milik Pribadi	aktif	layak
14	DESIH	LEMBANG SARI RT 017 RW 005	kurang dari UMKM	karyawan	Kontrak	aktif	layak
15	SITI KHOLIPAH	DUSUN LEMBANG SARI RT 018 RW 005	kurang dari UMKM	karyawan	Kontrak	aktif	layak
16	JULAHA	DUSUN CIBARENGKOK RT 010 RW 003	kurang dari UMKM	karyawan	Milik Pribadi	aktif	layak
17	CARSIH	LEMBANG SARI RT 017 RW 005	lebih dari UMKM	wiraswasta	Milik Pribadi	aktif	tidak layak
18	LULIS	DUSUN KRAJAN RT 007 RW 002	lebih dari UMKM	wiraswasta	Milik Pribadi	aktif	tidak layak
19	HOTIMAH	DUSUN CIBARENGKOK RT 011 RW 003	kurang dari UMKM	wiraswasta	Kontrak	aktif	layak
20	CICI HARTININGSIH	DUSUN KRAJAN RT 007 RW 002	kurang dari UMKM	wiraswasta	Kontrak	aktif	layak
21	RONI	DUSUN WATES RT 015 RW 004	kurang dari UMKM	wiraswasta	Milik Pribadi	aktif	layak
22	RINA MARDIANA	DUSUN KRAJAN RT 007 RW 002	kurang dari UMKM	karyawan	Milik Pribadi	aktif	layak
23	NENENG KURNIAWATI	KAMPUNG MUNGKUR JAYA RT 012 RW 003	kurang dari UMKM	karyawan	Kontrak	aktif	layak
24	ENTIN MASKINI	DUSUN BABAKAN SAPI RT 008 RW 002	lebih dari UMKM	wiraswasta	Milik Pribadi	aktif	tidak layak
25	R HERYANA	DUSUN KRAJAN RT 006 RW 002	lebih dari UMKM	wiraswasta	Milik Pribadi	aktif	tidak layak
26	SITI ARSAH	CIBARENGKOK RT 009 RW 003	kurang dari UMKM	karyawan	Milik Pribadi	aktif	layak
27	CARNESIH	KAMPUNG BABAKAN SUMUR RT 016 RW 011	kurang dari UMKM	karyawan	Milik Pribadi	aktif	layak
28	IAH MASLIJAH	KRAJAN 2 RT 007 RW 002	lebih dari UMKM	wiraswasta	Milik Pribadi	aktif	tidak layak
29	JAYAHAN	GANDASOLI RT 018 RW 013	lebih dari UMKM	wiraswasta	Milik Pribadi	aktif	tidak layak
30	SITI	DUSUN KRAJAN I RT 005 RW 001	kurang dari UMKM	wiraswasta	Milik Pribadi	aktif	layak
31	RESTI	DUSUN KRAJAN RT 007 RW 002	kurang dari UMKM	karyawan	Milik Pribadi	aktif	layak
32	IDA CAHYATI	DUSUN KRAJAN RT 007 RW 002	kurang dari UMKM	karyawan	Kontrak	aktif	layak
33	WARTO	DUSUN KRAJAN RT 007 RW 002	kurang dari UMKM	karyawan	Kontrak	aktif	layak
34	NASHI	KAMPUNG LEMBANG SARI RT 017 RW 005	kurang dari UMKM	wiraswasta	Kontrak	aktif	layak
35	EDAH	DUSUN MUNGKUR JAYA RT 012 RW 003	lebih dari UMKM	wiraswasta	Milik Pribadi	aktif	tidak layak
36	SUHAETI	DUSUN CIBARENGKOK RT 009 RW 003	lebih dari UMKM	wiraswasta	Milik Pribadi	aktif	tidak layak
37	MARYAM	KRAJAN I RT 003 RW 001	kurang dari UMKM	wiraswasta	Milik Pribadi	aktif	layak
38	IKEM	DUSUN KRAJAN RT 006 RW 002	kurang dari UMKM	wiraswasta	Milik Pribadi	aktif	layak
39	SURYATNI	DUSUN BABAKAN SAPI RT 008 RW 002	kurang dari UMKM	karyawan	Milik Pribadi	aktif	layak
40	TUTI HARVATI	KRAJAN I RT 002 RW 001	kurang dari UMKM	wiraswasta	Milik Pribadi	aktif	layak
41	DIKIK ROMAEINI	DUSUN CIBARENGKOK RT 010 RW 003	kurang dari UMKM	wiraswasta	Kontrak	aktif	layak
42	NYAI CARNESIH	DUSUN BABAKAN SAPI RT 008 RW 002	kurang dari UMKM	karyawan	Milik Pribadi	aktif	tidak layak
43	KARNI	KAMPUNG BABAKAN RESIK RT 001 RW 001	kurang dari UMKM	karyawan	Milik Pribadi	aktif	tidak layak
44	ATEM	DUSUN CIBARENGKOK RT 010 RW 003	kurang dari UMKM	karyawan	Milik Pribadi	aktif	layak
45	ATIN SUPRIATINI	DUSUN KRAJAN RT 007 RW 002	kurang dari UMKM	karyawan	Milik Pribadi	aktif	tidak layak
46	ILAH F	KRAJAN 2 RT 007 RW 002	kurang dari UMKM	karyawan	Milik Pribadi	aktif	layak
47	KANEM	CIBARENGKOK RT 010 RW 003	lebih dari UMKM	wiraswasta	Milik Pribadi	aktif	tidak layak
48	YUFUN MARYATI	DUSUN BABAKAN SAPI RT 008 RW 002	lebih dari UMKM	wiraswasta	Milik Pribadi	aktif	tidak layak
49	IAENI	DUSUN KRAJAN RT 007 RW 002	kurang dari UMKM	wiraswasta	Milik Pribadi	aktif	layak
50	SUKAESIH	DUSUN WATES RT 015 RW 004	kurang dari UMKM	wiraswasta	Kontrak	aktif	layak
51	MUSLIMIN	KRAJAN I RT 003 RW 001	kurang dari UMKM	wiraswasta	Kontrak	aktif	layak
52	IUMI	LEMBANG SARI RT 017 RW 011	kurang dari UMKM	wiraswasta	Kontrak	aktif	layak
53	ENCHI	LEMBANG SARI RT 017 RW 011	kurang dari UMKM	karyawan	Kontrak	aktif	layak
54	RATINI	KAMPUNG BABAKAN SUMUR RT 016 RW 011	kurang dari UMKM	karyawan	Milik Pribadi	aktif	tidak layak
55	AAAT R	KRAJAN RT 007 RW 011	kurang dari UMKM	karyawan	Milik Pribadi	aktif	tidak layak
56	TARISH	CIGADING RT 018 RW 013	kurang dari UMKM	karyawan	Milik Pribadi	aktif	layak
57	HIS	MUNGKUR JAYA RT 012 RW 003	kurang dari UMKM	karyawan	Kontrak	aktif	layak
58	EMPAT P	KRAJAN 2 RT 006 RW 002	lebih dari UMKM	wiraswasta	Milik Pribadi	aktif	tidak layak

Gambar 2. Potongan Dataset

Pengolahan Data

Dalam pengujian ini menggunakan *rapid miner* dengan *operator 10-fold cross-validation* untuk mendapatkan hasil *accuracy* yang tinggi pada setiap algoritma yang diuji menggunakan *dataset*. Berikut adalah proses Preprocessing.



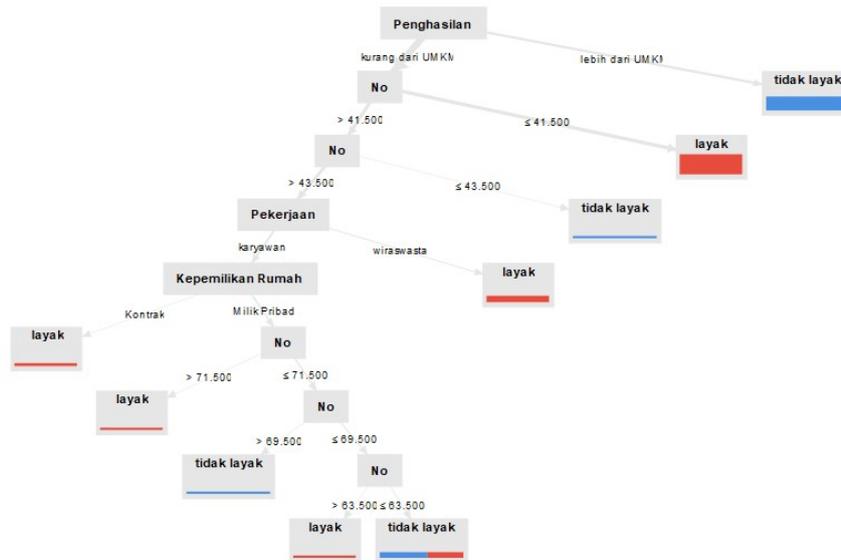
Pengukuran Penelitian

Berdasarkan hasil olah data melalui aplikasi Rapid Miner, maka diperoleh *Confusion Matrix* guna mengukur tingkat akurasi, dari algoritma C4.5, yaitu sebesar **100.00%**, *Precision* sebesar **100.00%**, dan *Sensitivity* atau *Recall* sebesar **100.00%**, seperti yang dapat kita lihat pada gambar dibawah ini.

accuracy: 100.00%

	true tidak layak	true layak	class precision
pred. tidak layak	2	0	100.00%
pred. layak	0	6	100.00%
class recall	100.00%	100.00%	

Berdasarkan data yang disajikan pada tabel di atas, akan menghasilkan sebuah model pohon keputusan seperti berikut :



Analisa Hasil Penelitian Data

Penjelasan dari model pohon keputusan yang dihasilkan :

Melalui data dari pohon keputusan yang menjadi atribut akar adalah penghasilan, dimana memiliki dua kelompok kelas (*range*) antara lain, kurang dari UMKM dengan lebih dari UMKM. Dari dua kelompok kelas tersebut kelompok kelas penghasilan kurang dari UMKM sudah mengklasifikasi kasus kedalam keputusan layak mendapatkan bantuan sosial, sedangkan kelompok kelas penghasilan lebih dari UMKM perlu dilakukan perhitungan untuk menentukan cabang selanjutnya, begitu juga setiap penjelasan node di bawahnya.

Berikut adalah *rule* yang berhasil di dapatkan dari pohon keputusan :

Tree

```

Penghasilan = kurang dari UMKM
|   No > 41
|   |   No > 43.500
|   |   |   Pekerjaan = karyawan
|   |   |   |   No > 71.500: layak {tidak layak=0, layak=4}
|   |   |   |   No ≤ 71.500
|   |   |   |   |   No > 69.500: tidak layak {tidak layak=2, layak=0}
|   |   |   |   |   No ≤ 69.500
|   |   |   |   |   |   Kepemilikan Rumah = Kontrak: layak {tidak layak=0, layak=3}
|   |   |   |   |   |   Kepemilikan Rumah = Milik Pribadi
|   |   |   |   |   |   |   No > 63.500: layak {tidak layak=0, layak=2}
|   |   |   |   |   |   |   No ≤ 63.500: tidak layak {tidak layak=3, layak=3}
|   |   |   |   |   |   |   Pekerjaan = wirawasta: layak {tidak layak=0, layak=9}
|   |   |   |   |   |   |   No ≤ 43.500: tidak layak {tidak layak=2, layak=0}
|   |   |   |   |   |   |   No ≤ 41: layak {tidak layak=0, layak=26}
Penghasilan = lebih dari UMKM: tidak layak {tidak layak=19, layak=0}

```

Tabel hasil pengolahan data dengan menggunakan metode Algoritma C4.5 dapat dilihat pada tabel dibawah ini :

Metode	Accuracy	Precision	Sensitivity atau Recall
Algoritma C4.5	100.00%	100.00%	100.00%

KESIMPULAN DAN SARAN

Dari hasil pengujian digunakan tes Algoritma C4.5 menghasilkan nilai *Accuracy* sebesar **100.00**, nilai untuk *Precision* sebesar **100.00**, dan nilai untuk *Sensitivity* atau *Recall* sebesar **100.00**. Dengan demikian Algoritma C4.5 merupakan algoritma dan teknik terbaik untuk Memprediksi Masyarakat dalam menerima bantuan sosial.

Saran untuk penelitian selanjutnya adalah dengan dilakukan sebuah komparasi dari beberapa Algoritma klasifikasi *data mining* yang lain, misalnya Algoritma C4.5 dengan Algoritma *K-Means* atau antara Algoritma C4.5, Algoritma *K-Means* dengan Algoritma *Naïve Bayes*, sehingga dapat diketahui Algoritma yang terbaik dalam Memprediksi Masyarakat Dalam Memperoleh Bantuan Sosial.

DAFTAR PUSTAKA

- [1] R. Fadilah and F. A. Siregar, "Bantuan Sosial Sembako dan Bantuan Sosial Tunai," 2021.
- [2] A. Septima Harefa, "Prosedur Penyaluran Bantuan Dana Jaminan Sosial Bagi Penyandang Cacat Berat Pada Dinas Sosial Kota Padang BAB I TA," Universitas Andalas, 2017.
- [3] M. Taufany Firmansyah and Rusito, "IMPLEMENTASI METODE DECISION TREE DAN ALGORITMA C4.5 UNTUK KLASIFIKASI DATA NASABAH BANK," 2016.
- [4] S. Hendrian, "Algoritma Klasifikasi Data Mining Untuk Memprediksi Siswa Dalam Memperoleh Bantuan Dana Pendidikan," *Faktor Exacta*, vol. 11, no. 3, Oct. 2018, doi: 10.30998/faktorexacta.v11i3.2777.
- [5] T. S. Mian and F. Ghabban, "Competitive Advantage: A Study of Saudi SMEs to Adopt Data Mining for Effective Decision Making," *Journal of Data Analysis and Information Processing*, vol. 10, no. 03, pp. 155–169, 2022, doi: 10.4236/jdaip.2022.103010.
- [6] I. Ba'abbad, T. Althubiti, A. Alharbi, K. Alfarsi, and S. Rasheed, "A Short Review of Classification Algorithms Accuracy for Data Prediction in Data Mining Applications," *Journal of Data Analysis and Information Processing*, vol. 09, no. 03, pp. 162–174, 2021, doi: 10.4236/jdaip.2021.93011.
- [7] N. Azwanti, "ALGORITMA C4.5 UNTUK MEMPREDIKSI MAHASISWA YANG MENGULANG MATA KULIAH (STUDI KASUS DI AMIK LABUHAN BATU)," *Jurnal SIMETRIS*, vol. 9, no. 1, 2018.
- [8] P. B. N. Setio, D. R. S. Saputro, and B. Winarno, "PRISMA, Prosiding Seminar Nasional Matematika Klasifikasi dengan Pohon Keputusan Berbasis Algoritme C4.5," vol. 3, pp. 64–71, 2020, [Online]. Available: <https://journal.unnes.ac.id/sju/index.php/prisma/>
- [9] R. Pratiwi, M. N. Hayati, and S. Prangga, "BAREKENG: Jurnal Ilmu Matematika dan Terapan PERBANDINGAN KLASIFIKASI ALGORITMA C5.0 DENGAN CLASSIFICATION AND REGRESSION TREE (STUDI KASUS: DATA SOSIAL KEPALA KELUARGA MASYARAKAT DESA TELUK BARU KECAMATAN MUARA ANCALONG TAHUN 2019) Comparison of C5.0 Algorithm Classification with Classification and Regression Tree (Case Study: Social Data of Family Head of Teluk Baru Village, Muara Ancalong District in 2019)," vol. 14, no. 2, 2020, doi: 10.30598/barekengvol14iss2pp267-278.
- [10] N. Azwanti, "ANALISA ALGORITMA C4.5 UNTUK MEMPREDIKSI PENJUALAN MOTOR PADA PT. CAPELLA DINAMIK NUSANTARA CABANG MUKA KUNING," 2018.
- [11] A. Abdulhafedh, "Comparison between Common Statistical Modeling Techniques Used in Research, Including: Discriminant Analysis vs Logistic Regression, Ridge Regression vs LASSO, and Decision Tree vs Random Forest," *OALib*, vol. 09, no. 02, pp. 1–19, 2022, doi: 10.4236/oalib.1108414.

- [12] S. Sekolah, T. Manajemen Informatika, D. Komputer, and W. Utama, "PENELITIAN KOMPARASI ALGORITMA KLASIFIKASI DALAM MENENTUKAN WEBSITE PALSU," vol. 1, no. 1, pp. 2598–294, 2017, [Online]. Available: <http://idsirtii.or.id>

NOMENKLATUR

S	: Himpunan Kasus
A	: Atribut
n	: Jumlah partisi S
p_i	: Proporsi dari S_i terhadap S
$ S_i $: Jumlah Kasus pada partisi ke i
$ S $: Jumlah Kasus dalam S
TP	: <i>True Positive</i>
TN	: <i>True Negatif</i>
FP	: <i>False Positive</i>
FN	: <i>False Negative</i>