

Klasifikasi Algoritma C4.5 dalam Menentukan Penerima Bantuan Covid-19 (Studi Kasus: Desa di Karawang)

Siti Badriah¹, Muhammad Fajar Estu Nugroho², Nurlana Sanjaya³, Imas Rismawati⁴,
Betha Nurina Sari⁵, Chaerur Rozikin⁶

^{1,2}Teknik Informatika, Ilmu Komputer, ³Universitas Singaperbangsa Karawang,
¹siti.badriah17201@student.unsika.ac.id, ²fajar.estu17147@student.unsika.ac.id,
³nurlana.sanjaya17167@student.unsika.ac.id, ⁴imas.16113@student.unsika.ac.id,
⁵betha.nurina@staff.unsika.ac.id, ⁶chaerur.rozikin@staff.unsika.ac.id

Abstrak

Covid-19 menyebabkan masyarakat Indonesia mengalami penurunan dalam pendapatan penghasilan. Dalam menanggulangi masalah penurunan pendapatan tersebut pemerintah memberikan bantuan bagi masyarakat akibat terdampak *Covid-19*. Berhubungan dengan adanya bantuan *Covid-19* tersebut, staf pemerintah desa kesulitan dalam melakukan proses seleksi penerima bantuan dan adanya ketidaktepatan dalam penerima bantuan *Covid-19*. Tujuan penelitian ini adalah membuat klasifikasi penerima bantuan sosial *Covid-19* sebagai bahan acuan dalam kebijakan penentuan penerima bantuan sosial. Sehingga dilakukan penelitian untuk mempermudah dalam penentuan dan ketepatan dalam bantuan *Covid-19*. Algoritma klasifikasi yang digunakan yaitu algoritma C4.5 dengan bahasa pemrograman php. Data yang digunakan terdapat 430 data dengan 5 atribut yaitu jumlah anggota keluarga, pekerjaan, penghasilan sebelum *Covid-19*, penghasilan setelah *Covid-19* dan keputusan (data per bulan April 2020). Hasil dari penelitian ini didapat *rule* dari algoritma C4.5 dengan dengan nilai akurasi yang terbaik terdapat pada perbandingan 90% *data training* dan 10% *data testing* dengan nilai akurasi sebesar 79,54%.

Kata Kunci: Klasifikasi, Bantuan *Covid-19*, Algoritma C4.5, *Confusion Matrix*.

1. Pendahuluan

Corona Virus jenis virus baru yang muncul pada akhir tahun 2019 yang kemudian disebut dengan *Corona Virus Disease (COVID-19)*. Menurut *World Health Organization (WHO)*, *covid-19* merupakan penyakit menular yang disebabkan oleh jenis corona virus yang dapat menyebabkan infeksi saluran nafas pada manusia. *Covid-19* telah menyebar di berbagai dunia, negara Indonesia juga termasuk salah satu diantaranya. Pemerintah Indonesia telah menetapkan Pembatasan Sosial Berskala Besar (PSBB) pada awal tahun 2020 untuk menimalisir jumlah masyarakat yang terjangkit *Covid-19*.

Akibat dari PSBB *Covid-19*, masyarakat Indonesia mengalami penurunan dalam pendapatan penghasilannya. Menurut CNN Indonesia (2020) mengenai *Survei Price Waterhouse Cooper (PWC)* mengungkapkan pandemi *Covid-19* telah menekan pendapatan rumah tangga masyarakat Indonesia. Tercatat, sebanyak 65 persen masyarakat Indonesia mengaku pendapatannya berkurang akibat pandemi *Covid-19* dan 31 persen responden mengaku masih memiliki pendapatan stabil. Sedangkan, hanya 4 persen dari responden yang mengaku pendapatannya bertambah selama pandemi berlangsung. Persentase masyarakat yang terdampak negatif finansialnya akibat pandemi di Indonesia lebih besar dibandingkan rata-rata global (CNN

Indonesia, 2020).

Dalam menanggulangi masalah penurunan pendapatan tersebut pemerintah memberikan bantuan bagi masyarakat akibat terdampak *Covid-19*. Berbagai bantuan dari pemerintah diberikan ke berbagai kalangan masyarakat, mulai dari pengangguran, karyawan swasta, UMKM, aparatur sipil Negara, siswa dan mahasiswa. Bantuan yang diberikan oleh pemerintah selama pandemi yaitu bantuan sembako dan uang bernilai 300.000 per bulan. Kabupaten Karawang juga termasuk ke dalam penerima saluran bantuan dari pemerintah. Berhubungan dengan adanya bantuan *Covid-19* tersebut, staf pemerintah desa kesulitan dalam melakukan proses seleksi penerima bantuan dan adanya ketidaktepatan dalam penerima bantuan *Covid-19*. Tujuan penelitian ini adalah membuat klasifikasi penerima bantuan sosial *Covid-19* sebagai bahan acuan dalam kebijakan penentuan penerima bantuan sosial.

Egi Badar Sambani dan Fitri Nuraeni (2017) telah melakukan penelitian mengenai klasifikasi pola penjurusan siswa di Sekolah Menengah Kejuruan (SMK) di kota Tasikmalaya dengan menggunakan algoritma C4.5. Ringga Sentagi Asa, Sarjon Defit, dan Jufriadif Na'am (2019) melakukan penelitian klasifikasi pola penyaluran zakat di Baznas Kabupaten Agam dengan menggunakan

algoritma C4.5. Hariati, Masna Wati, dan Bambang Cahyono (2018) juga menerapkan Algoritma C4.5 untuk penentuan penerima program bantuan pemerintah daerah di kabupaten Kutai Kartanegara.

Berdasarkan latar belakang di atas, maka akan dilakukan penelitian mengenai klasifikasi penerima bantuan Covid-19 dengan menggunakan algoritma C4.5.

2. Tinjauan Pustaka

2.1. Klasifikasi

Klasifikasi merupakan proses untuk menempatkan suatu objek ke salah satu kategori yang sudah ditentukan sebelumnya (Elisa, 2018). Klasifikasi termasuk ke *supervised learning* dan membutuhkan data latih untuk membangun suatu model klasifikasi (Ermawati, 2019). Model atau metode yang terdapat dalam jenis penyelesaian klasifikasi yaitu pohon keputusan, *naïve bayes*, Jaringan syaraf tiruan, analisis statistik, algoritma genetik, *rough set*, *K-nearest neighbor*, algoritma C4.5, metode berbasis aturan, *memory based reasoning*, *support vector machine* (Asa, 2019). Algoritma klasifikasi yang digunakan pada penelitian ini menggunakan klasifikasi C4.5.

2.2. Algoritma C4.5

Algoritma C4.5 merupakan metode klasifikasi dimana proses klasifikasi dimodelkan dengan menggunakan sekumpulan keputusan yang disusun dalam bentuk pohon struktur (Aggarwal, 2015). Algoritma C4.5 pengembangan dari algoritma ID3, sehingga prinsip dasar kerjanya sama dengan algoritma ID3 (Novianti et al., 2016). Pengembangan pada algoritma C4.5 dapat mengatasi *missing value*, *continue data* dan *pruning* (Rismayanti, 2016). Dalam menentukan pohon keputusan pada algoritma C4.5 terdapat 4 rumus sebagai berikut (Saifur Rohman Cholil & Ardianita, 2020):

Menghitung nilai *entropy* dengan melakukan persamaan:

$$Entropy(S) = -\sum p_i * \log_2 p_i \quad n \quad i=1 \quad (1)$$

Keterangan:

S = jumlah kasus

n = total himpunan S

p_i = proporsi ke-i

Menghitung nilai *Gain* dengan melakukan persamaan:

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

Keterangan:

S : jumlah kasus

A : karakter

n : total A

$|S_i|$: jumlah kasus pada partisi ke i

$|S|$: Jumlah kasus dalam S

Selanjutnya nilai *Split Info* dihitung dengan persamaan:

$$SplitInfo(S,A) = -\sum \frac{S_i}{S} \log_2 \frac{S_i}{S} \quad n \quad i=1 \quad (3)$$

Keterangan:

S = jumlah kasus

A = karakter

S_i = jumlah sampel karakter i

Yang menentukan sebuah atribut dapat dijadikan akar maupun cabang suatu pohon keputusan didapat dari nilai *Gain Ratio* dengan persamaan:

$$GainRatio(S,A) = \frac{Gain(S,A)}{SplitInfo(S,A)} \quad (4)$$

Keterangan:

Gain(S,A) = info gain pada karakter A

SplitInfo(S,A) = split info pada karakter A

2.3. Confusion Matrix

Confusion Matrix merupakan pengukuran kinerja untuk masalah klasifikasi pembelajaran mesin di mana keluaran dapat berupa dua kelas atau lebih. Hasil matriks dari prediksi dibandingkan dengan kelas yang asli dari masukan yang berisi nilai aktual dan prediksi pada klasifikasi (Dwi Kinasih Widiyati & Pakpahan, 2018). Berikut adalah Tabel dengan 4 kombinasi berbeda dari nilai prediksi dan nilai aktual (Sarang Narkhede, 2018).

Tabel 1 *Confusion matrix*

Nilai Prediksi	Nilai Aktual	
	Benar	Salah
Benar	BB	BS
Salah	SB	SS

Keterangan:

Benar Benar (BB) : Data benar terdeteksi benar.

Salah Benar (SB) : Data salah terdeteksi benar.

Benar Salah (BS) : Data benar terdeteksi salah.

Salah Salah (SS) : Data salah terdeteksi salah.

Berdasarkan nilai BB, SB, BS, dan SS maka akan diperoleh nilai akurasi dengan persamaan berikut:

$$Akurasi = \frac{(BB+SS)}{(BB+SS+SB+BS)} * 100 \quad (5)$$

Nilai akurasi menggambarkan seberapa akurat sistem dapat mengklasifikasikan data secara benar. Sebagai perbandingan kasus yang diidentifikasi benar dengan jumlah kasus (Waluyo, 2017).

2.4. Penelitian yang relevan

Suatu penelitian telah dilakukan oleh Nasrul, (2020) membuat sistem pendukung keputusan penerimaan bantuan dana koperasi desa menggunakan algoritma ID3 studi kasus di desa Galang Suka Kecamatan Galang, hasil dari penelitian ini menghasilkan *rule* dimana node pertama dengan hasil nilai gain tertinggi berada pada pembukuan dengan nilai 0,40406 dan didapat pola ID3. Zaman (2016) melakukan Penerapan algoritma C4.5 untuk menentukan kelayakan penerima bantuan rehabilitas sosial rumah tidak layak huni studi kasus di pemerintahan kabupaten Solok selatan, hasil menunjukkan pola pohon keputusan yaitu node pertama Lantai rumah dan node selanjutnya Sarana air bersih dan MCK . Hendrian

(2018) melakukan penerapan algoritma klasifikasi *data mining* untuk memprediksi siswa dalam memperoleh bantuan dana pendidikan, hasil menunjukkan nilai akurasi dari hasil pengujian *Cross Validation*, *Confusion Matrix* dan Kurva ROC sebesar 98,80%, nilai *precision* sebesar 98,02% dan nilai *recall* sebesar 99,00%. Dengan demikian algoritma C4.5 teknik terbaik dalam prediksi siswa dalam memperoleh bantuan

Penelitian selanjutnya Wintana et al (2019) mengenai klasifikasi penentuan penerima manfaat Program Keluarga Harapan (PKH) menggunakan algoritma C5.0 studi kasus di desa Sukamaju, hasil menunjukkan nilai *entropy* total 0,958 dan pohon keputusan yang didapat dengan *root* utama dari kepemilikan anak sekolah, lalu dilanjut dengan jenis dinding rumah, kepemilikan balita dan terakhir variabel pekerjaan. Eka Fitriani (2020) melakukan penelitian mengenai perbandingan algoritma C4.5 dan *naïve bayes* untuk menentukan kelayakan penerima bantuan program keluarga harapan, hasil menunjukkan algoritma C4.5 menghasilkan nilai *accuracy* 91,25% dan AUC 0,930, sedangkan untuk metode *naïve bayes* memiliki *accuracy* 87,11 dan AUC 0,923. Dengan demikian algoritma C4.5 metode yang lebih baik dalam memprediksi kelayakan warga dalam menerima bantuan Program Keluarga Harapan (PKH).

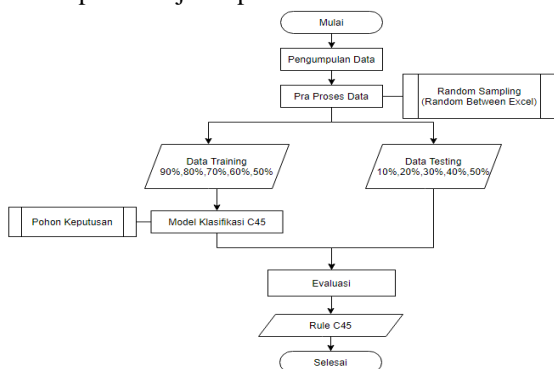
3. Metode Penelitian

3.1 Jenis Data

Jenis data yang digunakan pada penelitian ini adalah data kuantitatif dan data kualitatif. Data yang berbentuk angka ataupun bilangan yang dapat dihitung itu adalah data kuantitatif, sedangkan data yang berbentuk kalimat dan mempunyai unsur informasi adalah data kualitatif (Saifur Rohman Cholil & Ardianita, 2020). Data yang dikumpulkan adalah data primer karena data didapat dari narasumber pertama dari *staff* pegawai desa.

3.2 Tahapan Penelitian

Tahapan sistematis dilakukan pada penelitian ini seperti disajikan pada Gambar 1 berikut:



Gambar 1. Tahapan penelitian

4. Hasil Dan Pembahasan

4.1 Pengumpulan data

Data yang diperlukan yaitu data bantuan *Covid-19*. Pengumpulan data dilakukan secara langsung

melalui staf pemerintah desa yang mengelola data bantuan *Covid-19*. Data yang diperoleh sebanyak 430 data bantuan *Covid-19* per bulan April 2020. Adapun bentuk datanya pada Tabel 2.

Tabel 2. Data bantuan *Covid-19*

No	Nama	Jumlah Anggota Keluarga	Pekerjaan	Pendapatan Sebelum Covid-19	Pendapatan Setelah Covid-19	Keputusan
1	Wastini	4	Pedagang	100.000	50.000	Tidak
2	Sukardi	4	Wiraswasta	100.000	50.000	Tidak
3	Eka Sari	2	Buruh	100.000	50.000	Tidak
4	Awing	3	Buruh	100.000	50.000	Tidak
5	Riko	3	Buruh	100.000	50.000	Terima

Pada proses selanjutnya dilakukan eksplorasi data untuk mengetahui karakteristik data yang digunakan. Ekplorasi data atribut pada *dataset* dapat dilihat pada Tabel 3.

Tabel 3 Eksplorasi data

No.	Atribut	Tipe Data	Keterangan
1	Nama	Nominal	Nama masyarakat bantuan <i>Covid</i>
2	Jumlah anggota keluarga	Nominal	1 = 19 2 = 67 3 = 154 4 = 124 5 = 43 6 = 16 7 = 4 8 = 2 9 = 1
3	Pekerjaan	Nominal	Buruh = 260 Guru = 1 Karyawan swasta = 1 Mekanik = 1 Nelayan = 9 Pedagang = 17 Petani = 87 Supir = 2 Wartawan = 1 Wiraswasta = 51
4	Pendapatan Sebelum Covid-19	Nominal	50000 = 1 100000 = 387 150000 = 24 200000 = 6 300000 = 2 500000 = 9 1000000 = 1
5	Pendapatan Setelah Covid-19	Nominal	50000 = 390 75000 = 3 100000 = 25 120000 = 1 150000 = 1 200000 = 9 500000 = 1
6	Keputusan	Nominal	Terima = 134 Tidak = 296

4.2 Pra Proses Data

Pada pra proses data dilakukan dengan cara memilih atribut yang digunakan dan menghilangkan *missing value*. Atribut yang digunakan terdapat 5 atribut yaitu jumlah anggota keluarga, pekerjaan, pendapatan sebelum *covid*, pendapatan setelah *covid* dan keputusan. Pada data bantuan *Covid-19* yang berjumlah 430 data terdapat 3 data yang rangkap, sehingga data tersebut dihilangkan. Jadi total data yang digunakan berjumlah 427 data. Data 427 tersebut dilakukan pembagian *data training* 90%, 80%, 70%, 60%, 50% dan *data testing* 10%, 20%, 30%, 40%, 50%. Proses pengambilan *data training* dan *data testing* menggunakan fungsi *RANDOM BETWEEN* pada Microsoft Excel. Langkah selanjutnya menyimpan data ke dalam format .csv. Format .csv tersebut lalu dilakukan proses konversi ke dalam bentuk .json agar dapat dibaca oleh program PHP.

4.3 Model Klasifikasi C4.5

Penentuan bantuan *covid-19* pada proses algoritma C4.5 menggunakan bahasa pemrograman PHP. Implementasi dari hasil klasifikasi C4.5 membutuhkan hasil yang baik dilihat dari nilai akurasi yang didapat, sehingga dilakukan beberapa perbandingan pada *data training* dan *data testing*. Terdapat 5 perbandingan yang akan dilakukan yaitu 90% *data training* dan 10% *data testing*, 80% *data training* dan 20% *data testing*, 70% *data training* dan 30% *data testing*, 60% *data training* dan 40% *data testing*, 50% *data training* dan 50% *data testing*.

a. Hasil klasifikasi dengan 90 % *data training* dan 10% *data testing*

Hasil dari implementasi 90 % *data training* dan 10% *data testing* didapat *rule* algoritma C4.5 dalam bentuk pohon keputusan ditunjukkan pada gambar 2 dan nilai akurasi sebesar 79,54% ditunjukkan pada Tabel 4.

```

jumlah_anggota_keluarga = 1
pekerjaan = Pedagang : Tidak (1.0)
pekerjaan = Petani : Tidak (5.0)
jumlah_anggota_keluarga = 4
pendapatan_setelah_covid = 50000
pekerjaan = Pedagang
    pendapatan_sblm_covid = 100000 : Tidak (3.0/1.0)
    pendapatan_sblm_covid = 500000 : Terima (0.0)
    pekerjaan = Petani
        pendapatan_sblm_covid = 100000 : Tidak (20.0/6.0)
        pendapatan_sblm_covid = 500000 : Terima (0.0)
pendapatan_setelah_covid = 200000 : Terima (4.0)
    
```

Gambar 2. Pohon Keputusan dari 90 % *data training* dan 10% *data testing*

Tabel 4. Hasil nilai akurasi (a)

Nilai Prediksi	Nilai Aktual	
	Benar	Salah
Benar	3	7
Salah	32	2
Nilai Akurasi	79,54%	

b. Hasil klasifikasi dengan 80 % *data training* dan 20% *data testing*

Hasil dari implementasi 90 % *data training* dan 10% *data testing* didapat *rule* algoritma C4.5 dalam bentuk pohon keputusan ditunjukkan pada gambar 3 dan nilai akurasi sebesar 79,54% ditunjukkan pada Tabel 5.

```

jumlah_anggota_keluarga = 6
pekerjaan = Petani
    pendapatan_sblm_covid = 100000
        pendapatan_setelah_covid = 50000 : Terima (6.0/2.0)
        pendapatan_setelah_covid = 75000 : Tidak (0.0)
    pendapatan_sblm_covid = 150000 : Tidak (0.0)
pekerjaan = Buruh
    pendapatan_sblm_covid = 100000
        pendapatan_setelah_covid = 50000 : Tidak (4.0/1.0)
        pendapatan_setelah_covid = 75000 : Tidak (0.0)
    pendapatan_sblm_covid = 150000 : Tidak (1.0)
jumlah_anggota_keluarga = 3
pendapatan_setelah_covid = 50000
pekerjaan = Petani
    pendapatan_sblm_covid = 100000 : Tidak (20.0/5.0)
    pendapatan_sblm_covid = 150000 : Tidak (0.0)
pekerjaan = Buruh
    pendapatan_sblm_covid = 100000 : Tidak (82.0/32.0)
    pendapatan_sblm_covid = 150000 : Tidak (2.0/1.0)
pendapatan_setelah_covid = 75000 : Terima (1.0)
    
```

Gambar 3. Pohon Keputusan dari 80 % *data training* dan 20% *data testing*

Tabel 5. Hasil nilai akurasi (b)

Nilai Prediksi	Nilai Aktual	
	Benar	Salah
Benar	2	18
Salah	59	2
Nilai Akurasi	71,76%	

c. Hasil klasifikasi dengan 70 % *data training* dan 30% *data testing*

Hasil dari implementasi 70 % *data training* dan 30% *data testing* didapat *rule* algoritma C4.5 ditunjukkan oleh gambar 4 dalam bentuk pohon keputusan dan nilai akurasi sebesar 70,31% ditunjukkan pada Tabel 6.

```

jumlah_anggota_keluarga = 4
pekerjaan = Petani
    pendapatan_sblm_covid = 100000
        pendapatan_setelah_covid = 50000 : Tidak (18.0/6.0)
        pendapatan_setelah_covid = 75000 : Tidak (0.0)
    pendapatan_sblm_covid = 150000 : Tidak (2.0)
pekerjaan = Wiraswasta
    pendapatan_sblm_covid = 100000
        pendapatan_setelah_covid = 50000 : Tidak (9.0/3.0)
        pendapatan_setelah_covid = 75000 : Tidak (0.0)
    pendapatan_sblm_covid = 150000 : Terima (1.0)
jumlah_anggota_keluarga = 3
pendapatan_setelah_covid = 50000
pekerjaan = Petani
    pendapatan_sblm_covid = 100000 : Tidak (16.0/4.0)
    pendapatan_sblm_covid = 150000 : Tidak (0.0)
pekerjaan = Wiraswasta
    pendapatan_sblm_covid = 100000 : Tidak (6.0/2.0)
    pendapatan_sblm_covid = 150000 : Tidak (0.0)
pendapatan_setelah_covid = 75000
pekerjaan = Petani : Tidak (0.0)
pekerjaan = Wiraswasta : Terima (1.0)
    
```

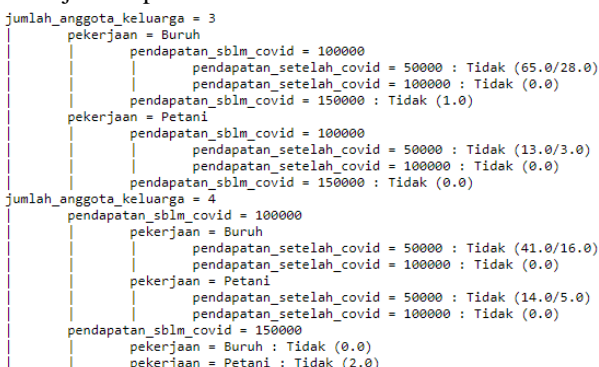
Gambar 4. Pohon Keputusan dari 70 % *data training* dan 30% *data testing*

Tabel 6. Hasil nilai akurasi (c)

Nilai Prediksi	Nilai Aktual	
	Benar	Salah
Benar	5	29
Salah	85	6
Nilai Akurasi	70,31%	

d. Hasil klasifikasi dengan 60 % *data training* dan 40% *data testing*

Hasil dari implementasi 60 % *data training* dan 40% *data testing* didapat *rule* algoritma C4.5 ditunjukkan pada gambar 5 dalam bentuk pohon keputusan dan nilai akurasi sebesar 70,76% ditunjukkan pada Tabel 7.



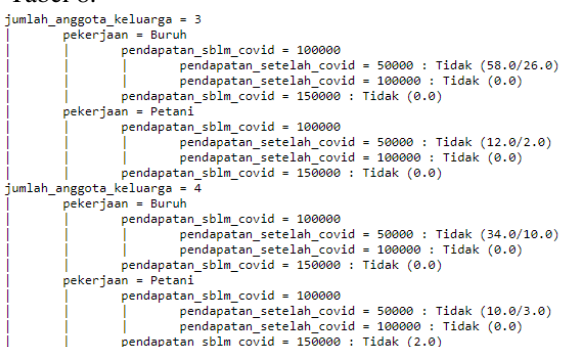
Gambar 5. Pohon Keputusan dari 60 % *data training* dan 40% *data testing*

Tabel 7. Hasil nilai akurasi (d)

Nilai Prediksi	Nilai Aktual	
	Benar	Salah
Benar	5	46
Salah	116	2
Nilai Akurasi	70,76%	

e. Hasil klasifikasi dengan 50 % *data training* dan 50% *data testing*

Hasil dari implementasi 50 % *data training* dan 50% *data testing* didapat *rule* algoritma C4.5 dalam bentuk pohon keputusan ditunjukkan pada gambar 6 dan nilai akurasi sebesar 66,66% ditunjukkan pada Tabel 8.



Gambar 6. Pohon Keputusan dari 50 % *data training* dan 50% *data testing*

Tabel 8. Hasil nilai akurasi (e)

Nilai Prediksi	Nilai Aktual	
	Benar	Salah
Benar	62	6
Salah	136	7
Nilai Akurasi	66,66%	

4.4 Pembahasan

Berdasarkan hasil klasifikasi di atas, maka pola pohon keputusan dan nilai akurasi yang didapat dengan melakukan 5 perbandingan yaitu 90% *data training* dan 10% *data testing* menghasilkan nilai akurasi sebesar 79,54%, 80% *data training* dan 20% *data testing* menghasilkan nilai akurasi sebesar 71,76%, 70% *data training* dan 30% *data testing* menghasilkan nilai akurasi sebesar 70,31%, 60% *data training* dan 40% *data testing* menghasilkan nilai akurasi sebesar 70,76%, 50% *data training* dan 50% *data testing* menghasilkan nilai akurasi sebesar 66,66%. Sehingga dalam penerapan klasifikasi penerima bantuan *covid-19* yang terbaik terdapat pada perbandingan 90% *data training* dan 10% *data testing* dengan nilai akurasi sebesar 79,54%. Adapun penjelasan dari pohon keputusan yang didapatkan disajikan pada Tabel 9.

Tabel 9 *Rule*

Id	Rule
1	If jumlah_anggota_keluarga = 1 And pekerjaan = Pedagang Then keputusan = Tidak
2	If jumlah_anggota_keluarga = 1 And pekerjaan = Petani Then keputusan = Tidak
3	If jumlah_anggota_keluarga = 4 And pendapatan_setelah_covid = 50000 And pekerjaan = Pedagang And pendapatan_sbmlm_covid = 100000 Then keputusan = Tidak
4	If jumlah_anggota_keluarga = 4 And pendapatan_setelah_covid = 50000 And pekerjaan = Pedagang And pendapatan_sbmlm_covid = 500000 Then keputusan = Terima
5	If jumlah_anggota_keluarga = 4 And pendapatan_setelah_covid = 50000 And pekerjaan = Petani And pendapatan_sbmlm_covid = 100000 Then keputusan = Tidak
6	If jumlah_anggota_keluarga = 4 And pendapatan_setelah_covid = 50000 And pekerjaan = Petani And pendapatan_sbmlm_covid = 500000 Then keputusan = Terima
7	If jumlah_anggota_keluarga = 4 And pendapatan_setelah_covid = 200000 Then keputusan = Terima

Pada Tabel 9 dapat diketahui bahwa *node* akar dari pohon keputusan adalah jumlah anggota keluarga. Dari *node* akar dapat dijelaskan bahwa masyarakat yang memiliki jumlah anggota keluarga 4 diterima sebagai penerima bantuan *Covid-19* jika pekerjaan pedagang dan petani, pendapatan sebelum covid sebesar 500000 dan pendapatan setelah covid 50000. Hal ini menunjukkan bantuan diberikan kepada masyarakat yang mengalami kerugian atau penurunan pendapatan yang drastis.

5. Kesimpulan dan Saran

5.1 Kesimpulan

Berdasarkan pembahasan di atas, maka dapat disimpulkan bahwa hasil dari penelitian ini didapat *rule* dari algoritma C4.5 dengan dengan nilai akurasi yang terbaik terdapat pada perbandingan 90% *data training* dan 10% *data testing* dengan nilai akurasi sebesar 79,54%. Pohon keputusan yang didapat terdapat atribut yang sangat berpengaruh yaitu atribut jumlah anggota keluarga yang dapat dijadikan sebagai bahan acuan diterima. Penelitian ini hanya menggunakan 5 atribut sebagai data untuk penentuan penerima bantuan sosial, mungkin penelitian selanjutnya bisa menambahkan atribut yang lain yang mendukung keputusan penentuan penerima bantuan sosial.

5.2 Saran

Saran atau rekomendasi selanjutnya bisa menggunakan algoritma atau metode lain seperti metode SMOTE untuk menangani data tidak seimbang.

Daftar Pustaka

Aggarwal, C. C. (2015). Data Mining. In *Data Mining*.

Asa, R. S. (2019). Identifikasi Penyaluran Zakat Menggunakan Algoritma C4.5 (Studi Kasus Di Baznas Kabupaten Agam). *Jurnal Sains Dan Informatika*, 5(1), 50. <https://doi.org/10.22216/jsi.v5i1.4048>

Cnn Indonesia. (2020). *Survei: Pandemi Corona Gerus Pendapat Masyarakata*. Cnnindonesia. <https://www.cnnindonesia.com/ekonomi/20200813165108-92-535427/survei-pandemi-corona-gerus-pendapat-masyarakata>

Dwi Kinasih Widiyati, M. W., & Pakpahan, H. S. (2018). Penerapan Algoritma Id3 Decision Tree Pada Penentuan Penerima Program Bantuan Pemerintah Daerah Di Kabupaten Kutai Kartanegara. *Jurti*, 2(2), 125–134.

Eka Fitriani. (2020). Perbandingan Algoritma C4.5 Dan Naïve Bayes Untuk Menentukan Kelayakan Penerima Bantuan Program Keluarga Harapan. *Sistemasi*, 9, 103–115.

Elisa, E. (2018). Prediksi Profit Pada Perusahaan Dengan Klasifikasi Algoritma C4.5. *Kumpulan Jurnal Ilmu Komputer (Klik)*, 05(02), 179–189.

Ermawati, E. (2019). Algoritma Klasifikasi C4.5 Berbasis Particle Swarm Optimization Untuk Prediksi Penerima Bantuan Pangan Non Tunai. *Sistemasi*, 8(September), 513–528.

Hendrian, S. (2018). Algoritma Klasifikasi Data Mining Untuk Memprediksi Siswa Dalam Memperoleh Bantuan Dana Pendidikan. *Faktor Exacta*, 11(3), 266–274.

Nasrul, A. (2020). Sistem Pendukung Keputusan Penerimaan Bantuan Dana Koperasi Desa

Menggunakan Algoritma Interative Dichotomiser 3 (Id3) (Studi Kasus : Desa Galang Suka Kecamatan Galang). *Terapan Informatika Nusantara*, 1(1), 17–25.

Novianti, B., Rismawan, T., & Bahri, S. (2016). Implementasi Data Mining Dengan Algoritma C4 . 5 Untuk Penjurusan Siswa (Studi Kasus : Sma Negeri 1 Pontianak). *Coding*, 04(3).

Pakpahan, H. S., Indar, F., Wati, M., Teknologi, J., & Mulawarman, U. (2018). *Penerapan Algoritma Cart Decision Tree Pada Penentuan Penerima Program Bantuan Pemerintah Daerah Kabupaten Kutai Kartanegara*. 2(1), 27–36.

Rismayanti. (2016). Implementasi Algoritma C4.5 Untuk Menentukan Penerima Beasiswa Di Stt Harapan Medan. *Media Infotama*, 12(2), 116–120.

Saifur Rohman Cholil, A. F. D., & Ardianita, T. (2020). Prediksi Penyakit Demam Berdarah Di Puskesmas Ngemplak Simongan Menggunakan Algoritma C4.5. *Sistemasi*, 9(September), 529–542.

Sambani, E. B., & Nuraeni, F. (2018). Penerapan Algoritma C4.5 Untuk Klasifikasi Pola Penjurusan Di Sekolah Menengah Kejuruan (Smk) Kota Tasikmalaya. *Csrid (Computer Science Research And Its Development Journal)*, 9(3), 144. <https://doi.org/10.22303/csrid.9.3.2017.144-152>

Sarang Narkhede. (2018). *Confusion Matrix*. Towardsdatascience. <https://towardsdatascience.com/understanding-confusion-matrix-a9ad42dcfd62>

Waluyo, S. H. (2017). *Klasifikasi Pemanfaat Program Beras Sejahtera (Rastra) Berdasarkan Tingkat Kemiskinan Dengan Menggunakan Algoritma Decision Tree C4 . 5 Berbasis Particle Swarm Optimization*. 7(2), 19–24.

Wati, M., Cahyono, B., Mulawarman, U., Mining, D., & Kartanegara, B. K. (2018). Penerapan Algoritma C4 . 5 Pada Penentuan Penerima Program Bantuan Pemerintah Daerah Di Kabupaten Kutai Kartanegara. *Jurti*, 2(2), 106–114.

Wintana, D., Hikmatulloh, N. I., & Purnama, Jajang Jaya, Rahmawati, A. (2019). Klasifikasi Penentuan Penerima Manfaat Program Keluarga Harapan (Pkh) Menggunakan Algoritma C5.0 (Studi Kasus: Desa Sukamaju, Kec.Kadudampit). *Kumpulan Jurnal Ilmu Komputer (Klik)*, 06(03), 254–264.

Zaman, K. (2016). Penerapan Data Mining Menggunakan Algoritma C4 . 5 Untuk Menentukan Kelayakan Penerima Bantuan Rehabilitas Sosial Rumah Tidak Layak Huni (Studi Kasus Di Pemerintahan Kabupaten Solok Selatan). *Komtek*, 3(2), 12–24.