

Implementasi Algoritma C.45 dalam Klasifikasi Kondisi Ekonomi Warga Kabupaten Boyolali

Andhi Prasetyo¹, Ari Wahyono²

^{1,2}Teknik Informatika, Fakultas Komunikasi dan Teknik Informatika, Universitas Boyolali

E-mail: ¹andhi.prasetyo.ti@gmail.com, ²namaku.ariwahyono@gmail.com

Corresponden Author: andhi.prasetyo.ti@gmail.com

Diterima Redaksi: 14 Juni 2024 Revisi Akhir: 14 September 2024 Diterbitkan Online: 22 September 2024

Abstrak – Dampak ekonomi setelah pandemi Covid-19 dirasakan berbagai negara. Pemerintah Indonesia melaksanakan pemulihan ekonomi dan memberikan bantuan sosial berdasarkan kondisi ekonomi, namun masih ada warga yang tidak mendapatkan bantuan padahal layak untuk dibantu. Pemerintah Desa Sumbang berusaha mengantisipasi dengan mencari pola yang menjadi indikator paling berpengaruh pada kondisi ekonomi untuk kelayakan pemberian bantuan sosial.

Penelitian dilakukan menggunakan metode klasifikasi yaitu algoritma C4.5 karena menghasilkan visualisasi pohon keputusan yang mudah dipahami. Data yang digunakan adalah data MCD Boyolali untuk Desa Sumbang Kecamatan Cepogo. Hasil penelitian dari 21 atribut, diperoleh 8 kriteria dengan bobot tertinggi sebagai indikator kondisi ekonomi. Hasil akurasi algoritma C4.5 mencapai 94,47%, lebih tinggi dari Naïve Bayes (93,28%) dan K-NN (91,70%) sehingga cocok untuk melakukan klasifikasi kondisi ekonomi warga Kabupaten Boyolali khususnya Desa Sumbang.

Kata Kunci — Algoritma C4.5, Data Mining, Kondisi Ekonomi, MCD Boyolali, Pohon Keputusan

Abstract – The economic impact of the Covid-19 pandemic was felt by various countries. The Indonesian government implemented economic recovery and provided social assistance based on economic conditions, but there are still residents who do not receive assistance even though they deserve help. The Sumbang Village Government is trying to anticipate by looking for patterns that are the most influential indicators of economic conditions for the eligibility of providing social assistance.

The research was carried out using the classification method, namely the C4.5 algorithm because it produces decision tree visualizations that are easy to understand. The data used is MCD Boyolali data for Sumbang Village, Cepogo District. The research results of 21 attributes obtained 8 criteria with the highest weight as indicators of economic conditions. The accuracy results of the C4.5 algorithm reached 94.47%, higher than Naïve Bayes (93.28%) and K-NN (91.70%) so it is suitable for classifying the economic conditions of residents of Boyolali Regency, especially Sumbang Village.

Keywords — C4.5 Algorithm, Data Mining, Decision Tree, Economic Conditions, MCD Boyolali



1. PENDAHULUAN

Wabah Covid-19 berdampak pada bidang sosial, politik maupun ekonomi. Pasca Wabah Covid-19, pemerintah melaksanakan kebijakan Pemulihan Ekonomi yaitu rangkaian kegiatan yang bertujuan mengurangi dampak negatif yang telah ditimbulkan dengan memberikan bantuan sosial [1]. Bantuan sosial (Bansos) merupakan program program dibebankan pada Anggaran Pendapatan dan Belanja Pemerintah. Bantuan sosial dapat diberikan pada perseorangan, keluarga maupun kelompok masyarakat namun tidak bersifat berkelanjutan dan bersifat selektif [2]. Bantuan sosial hanya diberikan kepada warga kurang mampu secara ekonomiyaitu tidak mampu memenuhi kebutuhan dasarnya seperti kebutuhan rohani, pangan, sandang, papan, dan kesehatan. [3]. Namun, sering kali masih ditemukan penerima bansos yang salah sasaran, tidak merata dalam penyaluran bansos, dan masih ditemukan data ganda [4]. Dalam rangka antisipasi, Pemerintah Desa Sumbang ingin menemukan indikator yang mempengaruhi kondisi ekonomi yang layak mendapatkan bantuan sehingga tidak salah sasaran. Metode yang tepat adalah dengan metode klasifikasi data mining [5] agar mendapatkan kesimpulan atau pola yang menjadi indikator bahwa seseorang dalam kondisi ekonomi yang layak dibantu. Kondisi ekonomi merupakan kemampuan

seseorang untuk memenuhi kebutuhan hidupnya [6]. Kemensos membaginya menjadi 7 kriteria [7] dan BPS (Badan Pusat Statistik) menjadi 14 kriteria [8].

Metode klasifikasi menggunakan Algoritma C4.5 yang menawarkan keuntungan seperti hasil yang mudah divisualisasikan, penghapusan fitur yang tidak relevan (noise), dan memiliki akurasi lebih baik berdasarkan komparasi C4.5 (77,91%), *Naïve Bayes* (76,59%), *K-NN* (63,54%) dalam analisis keberlangsungan pasien gagal jantung [9]. Dalam penelitian lain Algoritma C4.5 memiliki performa lebih baik daripada *Naïve Bayes* [10], [11], [12].

Penelitian ini berusaha mencari akurasi dari algoritma C4.5 dan indikator yang paling berpengaruh dalam penentuan kondisi ekonomi warga kabupaten Boyolali. Label yang diberikan yaitu “TIDAK LAYAK” (untuk status “MAMPU”) dan “LAYAK” (untuk status “RAWAN KEMISKINAN”, “MISKIN” dan “MISKIN ABSOLUT”) karena *MCD* Boyolali menjadi rujukan pemberian bantuan sosial [13]. Penelitian menggunakan sampel data *MCD* (*Monitoring Center for Development*) Boyolali yaitu data *MCD* Desa Sumbang Kecamatan Cepogo Kabupaten Boyolali Bulan April Tahun 2024 dengan 37 atribut, 4368 baris data dan 1266 data Kepala Keluarga. Dengan hasil yang diperoleh diharapkan Pemerintah Desa Sumbang dapat melaksanakan program pengentasan kemiskinan dan dapat mengurangi resiko bantuan yang salah sasaran.

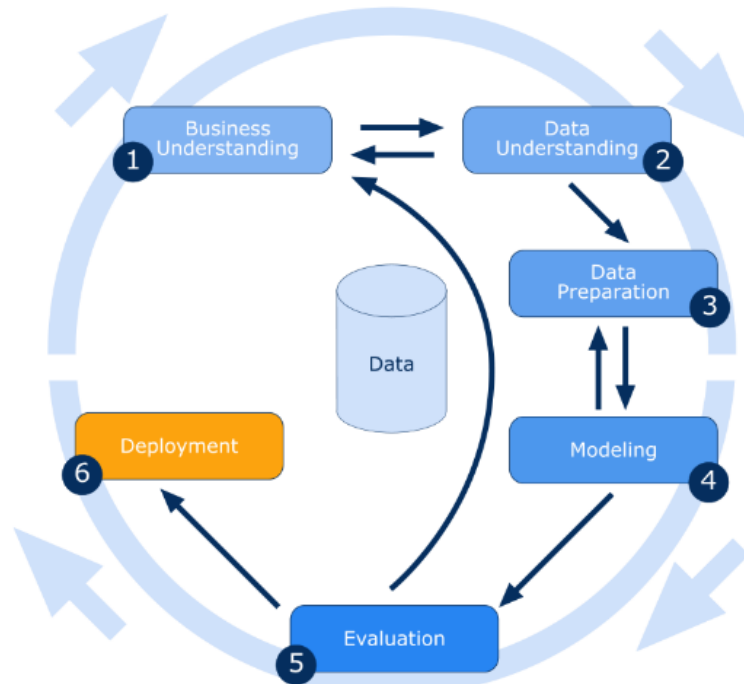
2. METODE PENELITIAN

Tahap awal dengan studi lapangan dan wawancara di wilayah penelitian, perumusan masalah, studi pustaka, tujuan penelitian dilanjutkan dengan pemrosesan data. Metode yang digunakan adalah metode standar dalam penelitian *Data Mining* yaitu *Cross Industry Standard Process for Data Mining* (*CRISP-DM*) [10], [14], [15], [16], [14], [17], [18]. Sedangkan untuk pengujiannya menggunakan *Confusion Matrix* sehingga secara langsung dapat melihat hasil akurasi [10].

2.1. Pengumpulan Data

Data yang menjadi basis dalam penelitian ini diperoleh dari Admin *MCD* Desa Sumbang yang berisi nama, nik, status perkawinan, usia, pendidikan, penghasilan, kepemilikan Tabungan, kondisi rumah, kepemilikan jaminan Kesehatan maupun akses informasi.

2.2. CRISP-DM



Gambar 1. Diagram CRISP-DM

2.2.1. *Business Understanding* (Pemahaman Bisnis)

Pemahaman terkait dengan masalah yang sedang terjadi yaitu tentang bantuan sosial dan tujuan yang ingin dicapai yaitu menemukan akurasi klasifikasi *Data Mining* dan indikator dengan bobot tertinggi yang berpengaruh dalam kondisi ekonomi warga.

2.2.2. *Data Understanding* (Pemahaman Data)

Pemahaman tentang setiap atribut data dari data *MCD* Boyolali. Pemahaman terkait atribut data dan nilai data perlu dilakukan agar memberikan gambaran besar dari data *MCD* Boyolali khususnya data *MCD* Boyolali untuk Desa Sumbang Kecamatan Cepogo.

2.2.3. *Data Preparation* (Persiapan Data)

Untuk menghasilkan data yang berkualitas, diperlukan proses *preprocessing* sebelum menggunakan data pada operasi utama dalam pengolahan data. Proses *preprocessing* dalam penelitian ini menggunakan *Microsoft Excel*.

2.2.3.1. *Data Cleaning* (Pembersihan Data)

Proses ini memastikan keakuratan, konsistensi, dan kegunaan data. Prosesnya yang dilakukan adalah mendeteksi kesalahan data atau data yang *corrupt* (rusak / tidak sempurna) dan memperbaiki atau menghapus data sesuai kebutuhan.

2.2.3.2. *Data Transformation* (Transformasi Data)

Digunakan dalam proses data mining untuk mengubah data ke dalam format yang sesuai melalui teknik *normalization* (normalisasi), *attribute selection*, dan *discretization*.

2.2.3.3. *Data Redundancy* (Reduksi Data)

Memilah data yang berukuran besar tentu sangat melelahkan, karena itu perlu dilakukan pengurangan data melalui penyederhanaan data, pemilihan atribut, penyesuaian nilai atribut, maupun pengurangan ukuran data [19].

2.2.4. *Modeling* (Pembuatan Model)

Dalam tahapan ini, data yang telah dibersihkan diimport ke dalam aplikasi *RapidMiner* dan mulai pembentukan model klasifikasi. Pemodelan ini memperhatikan kedalaman dan pemangkasan pohon keputusan.

2.2.5. *Evaluation* (Evaluasi)

Hasil yang telah diperoleh diuji dan dibandingkan dengan kondisi yang berbeda untuk menemukan performa terbaik sehingga layak untuk digunakan. Dalam evaluasi, hasil analisis klasifikasi dievaluasi berdasarkan keadaan sebenarnya.

2.2.6. *Deployment*

Tahap implementasi atas model yang sudah terbentuk dalam sebuah simulasi. Dalam penelitian menggunakan simulasi visual melalui *RapidMiner*. Visualisasi dengan pohon keputusan bertujuan agar presentasi dapat efektif dan mudah dipahami.

2.3. *Rapid miner*

RapidMiner adalah software yang berfungsi untuk analisis *data mining* dan merupakan aplikasi terbaik untuk *Data Mining* [5], [20].

2.4. *Algoritma C4.5*

Algoritma C4.5 merupakan salah satu algoritma klasifikasi untuk melakukan peramalan atau prediksi suatu data dengan hasil perupa pohon keputusan [21]. Algoritma C4.5 melalui beberapa tahapan sebagai berikut:

2.4.1. *Menghitung nilai entropy*

Perhitungan pertama adalah mengukur tingkat keragaman nilai yang ada pada sebuah kolom terhadap kriteria pada label hasil [22], [23].

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

Keterangan:

- S : Jumlah seluruh data
- n : Jumlah kumpulan data
- pi : Nilai atribut dibagi label hasil
- log2 : Logaritma biner / basis dua

Catatan: Pada saat nilai pada sebuah kolom dengan label identik, *entropy* =1, dan jika salah satu >0, nilai *entropy* = 0.

2.4.2. Menghitung nilai gain

Gain(S,A) merupakan bobot yang dihitung dari selisih *entropy* total dikurangi *entropy* atribut dikalikan jumlah pada kolom dibagi data label. [22].

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

Keterangan:

- S : Jumlah seluruh data
- A : Atribut atau kolom
- n : Jumlah partisi S
- |Si| : Jumlah kasus pada partisi ke-i
- |S| : Jumlah kasus dalam S

2.4.3. Menghitung SplitInfo

SplitInfo menyatakan informasi *entropy* dari nilai atribut terhadap data keseluruhan [22].

$$SplitInfo(S,A) = - \sum_{i=1}^n \frac{S_i}{S} * \log_2 \frac{S_i}{S} \dots\dots\dots (3)$$

Keterangan:

- S : Jumlah seluruh data
- A : Atribut atau kolom
- Si : Jumlah sampel untuk atribut A
- log2 : Logaritma biner / basis dua

2.4.4. Menghitung Gain Ratio

Gain Ratio digunakan untuk mengurangi bias atribut di banyak cabang [22].

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

Keterangan:

- S : Himpunan kasus
- A : Atribut
- Gain(S,A) : Gain A
- SplitInfo(S,A) : SplitInfo A

2.5. Confusion Matrix

Confusion Matrix adalah metode untuk mengevaluasi kinerja algoritma klasifikasi [24].

Tabel 1. Confusion Matrix

Aktual	Prediksi data	
	Negatif	Positif
Negatif	TN (<i>True Negative</i>)	FN (<i>False Negative</i>)
Positif	FP (<i>False Positive</i>)	TAPI (<i>True Positive</i>)

Keterangan:

- a. *True Negatif* (TN): bila data sebenarnya bernilai negatif, hasil prediksi menunjukkan negatif.
- b. *True Positif* (TP): bila data sebenarnya bernilai positif, hasil prediksi menunjukkan positif.
- c. *False Negatif* (FN): bila data sebenarnya bernilai positif, hasil prediksi menunjukkan negatif.
- d. *False Positif* (FP): bila data sebenarnya bernilai negatif, hasil prediksi menunjukkan positif.

2.5.1. Accuracy

Rasio prediksi dengan jumlah total prediksi untuk menilai seberapa baik hasil klasifikasi yang dilakukan oleh suatu model sesuai dengan nilai yang sebenarnya. [9].

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \dots\dots\dots (5)$$

2.5.2. Precision

Rasio prediksi benar dengan jumlah total prediksi positif untuk mengukur seberapa baik model dalam mengidentifikasi sampel positif yang sebenarnya. [9].

$$Precision = \frac{TP}{TP+FP} \times 100\% \dots\dots\dots (6)$$

2.5.3. Recall

Rasio prediksi dengan jumlah total prediksi untuk mengukur seberapa baik model dalam menemukan sampel yang sebenarnya positif [9].

$$Recall = \frac{TP}{TP+FN} \times 100\% \dots\dots\dots (7)$$

2.5.4. f-measure

Nilai f-measure didapatkan dengan memadukan nilai recall dan precision [9].

$$f - measure = \frac{2 \times recall \times precision}{recall+precision} \times 100\% \dots\dots\dots (8)$$

3. HASIL DAN PEMBAHASAN

Proses analisis diawali dengan persiapan alat bantu Data Mining yaitu RapidMiner, dilanjutkan dengan proses CRISP-DM terhadap data yang telah tersedia.

3.1. Persiapan Aplikasi Rapid Miner

Aplikasi RapidMiner awalnya dapat diunduh melalui www.rapidminer.com, atau <https://altair.com/altair-rapidminer>. Aplikasi RapidMiner saat ini telah diakuisisi oleh perusahaan yang bernama Altair Engineering pada bulan September tahun 2022. Setelah berhasil download, dilanjutkan instalasi dalam komputer.

3.2. Pengumpulan data

Setelah data diperoleh, data dipelajari struktur dan kondisinya agar menemukan cara yang paling efektif dalam persiapan data. Data yang diperoleh berupa data keluarga namun pengisian informasi secara menyeluruh hanya pada kepala keluarga saja.

No	Nama Warga	NIK	Status Hubungan dlm Keluarga	Status Perkawinan	Jenis kelamin	Tgl/bh/ lahir	Usia	Pendidikan terakhir yang ditamatkan	Pekerjaan	Penghasilan /bulan	Kepemilikan harta lancar	kemampuan konsumsi	rasio pengeluaran pangan	jenis konsumsi	kemampuan membeli pakaian	status tempat tinggal	luas lantai	Jenis lantai
1	EKO NUJR SETI	33090706	kepala Keluarga	Kawin	L	06/04/198	39	SMA	BURUH TETA	2.000.001-3.0	TABUNGAN ATAU BARANG	MAKAN TIGA KAL	70% UNTUK MAKAN	DENGAN LAU	LEBIH DARI SEK	RUMAH SENDIRI	=>BM PERSE	30% TANAH
2	TRIVYANI	33090359	Suami/Istri	Kawin	P	13/04/198	38	SMA	TIDAK BEKERJA/SEKOLAH									
3	ABYAN SHAKA	33090308	Anak	Belum Kawin	L	08/10/205	6	TIDAK TAMAT SD	TIDAK BEKERJA/SEKOLAH									
4	JEFFRI ARDIAN	33090301	kepala Keluarga	Kawin	L	01/02/199	27	D1/D2/D3/S1/S2	TWRASWASTI	3-3000.000	TABUNGAN ATAU BARANG	MAKAN TIGA KAL	30% UNTUK MAKAN	DENGAN LAL	LEBIH DARI SEK	RUMAH SENDIRI	=>BM PERSE	PERMANEN
5	RENI SAPUTRI	33221758	Suami/Istri	Kawin	P	18/06/199	25	D1/D2/D3/S1/S2	BURUH TETAP									
6	SRI SUNARMI	33090341	familia lain	Cerai Mati	P	01/02/196	63	SMA	PNS/TNI/POLRI									
7	JEFFRI ARDIAN	3309030102970004				01/02/199	27											
8	SLAMET	33090308	kepala Keluarga	Kawin	L	06/04/197	54	SMA	PETANI	2.000.001-3.0	TABUNGAN ATAU BARANG	MAKAN TIGA KAL	70% UNTUK MAKAN	DENGAN LAL	LEBIH DARI SEK	RUMAH SENDIRI	=>BM PERSE	30% TANAH
9	MUJINEM	33090370	Suami/Istri	Kawin	P	30/07/197	49	SMA	PETANI									
10	EDI LISTANTO	33090316	Anak	Belum Kawin	L	16/09/199	27	SMA	BURUH TETAP									
11	ANTON KURNI	33090326	Anak	Belum Kawin	L	26/12/200	21	SMP	BURUH TIDAK TETAP									
12	TRIVYANI APRIL	33090328	Anak	Belum Kawin	P	12/04/200	16	SD	TIDAK BEKERJA/SEKOLAH									
13	BURAT MARI	33090341	familia lain	Cerai Mati	P	01/07/194	88	TIDAK TAMAT SD	PETANI									
14	SLAMET BASU	33090320	kepala Keluarga	Kawin	L	20/05/196	56	SMA	PETANI	2.000.001-3.0	TABUNGAN ATAU BARANG	MAKAN TIGA KAL	70% UNTUK MAKAN	DENGAN LAL	LEBIH DARI SEK	RUMAH SENDIRI	=>BM PERSE	PERMANEN
15	MULYATI	33090370	Suami/Istri	Kawin	P	30/06/197	53	SMA	BURUH TIDAK TETAP									
16	TUKINEM	33090341	kepala Keluarga	Cerai Mati	P	01/07/193	88	SD	PETANI	600.000-900.	TABUNGAN ATAU BARANG	MAKAN TIGA KAL	70% UNTUK MAKAN	DENGAN LAL	LEBIH DARI SEK	RUMAH SENDIRI	=>BM PERSE	PERMANEN
17	ANDHYANTO	33011412	kepala Keluarga	Kawin	L	12/01/197	52	SMA	BURUH TIDAK TETAP	900.001-2.000	TABUNGAN ATAU BARANG	MAKAN TIGA KAL	70% UNTUK MAKAN	DENGAN LAL	LEBIH DARI SEK	RUMAH SENDIRI	=>BM PERSE	30% TANAH
18	NUJR AMINAH	33011413	Suami/Istri	Kawin	P	01/04/197	45	SMP	TIDAK BEKERJA/SEKOLAH									
19	AMBAR DWI	33011454	Anak	Belum Kawin	P	14/01/200	15	SD	TIDAK BEKERJA/SEKOLAH									
20	BAGAS OKTA	33011423	Anak	Belum Kawin	L	23/10/200	22	SMA	TIDAK BEKERJA/SEKOLAH									

Gambar 2. Data sebelum diproses (Bagian 1)

Gambar 3. Data sebelum diproses (Bagian 2)

3.3. Pemrosesan data dalam Microsoft excel

Melalui *Microsoft excel*, data disorting dan menemukan *noise* dengan cepat. hasil akhir dari pemrosesan ini adalah data penghapusan data non kepala keluarga dan dihasilkan data kepala keluarga dengan tambahan atribut jumlah tanggungan.

Gambar 4. Penghapusan noise dan transformasi data

3.4. Penghapusan kolom yang tidak diperlukan

Data yang sudah dihasilkan adalah 1266 baris data, dilakukan pemberian interval usia pada kolom usia dengan mengacu pada data *The World Health Organization (WHO)* [25]. Kemudian menghapus kolom-kolom yang tidak begitu berpengaruh dalam Analisa yaitu kolom “No”, “Nama”, “NIK”, “No KK” dll.

3.5. Pemilihan label dan penyeragaman istilah

Label yang digunakan diseragamkan menjadi “LAYAK” dan “TIDAK LAYAK”, kemudian dalam nilai atribut disederhanakan agar tidak mengganggu pada saat pembentukan pohon keputusan.

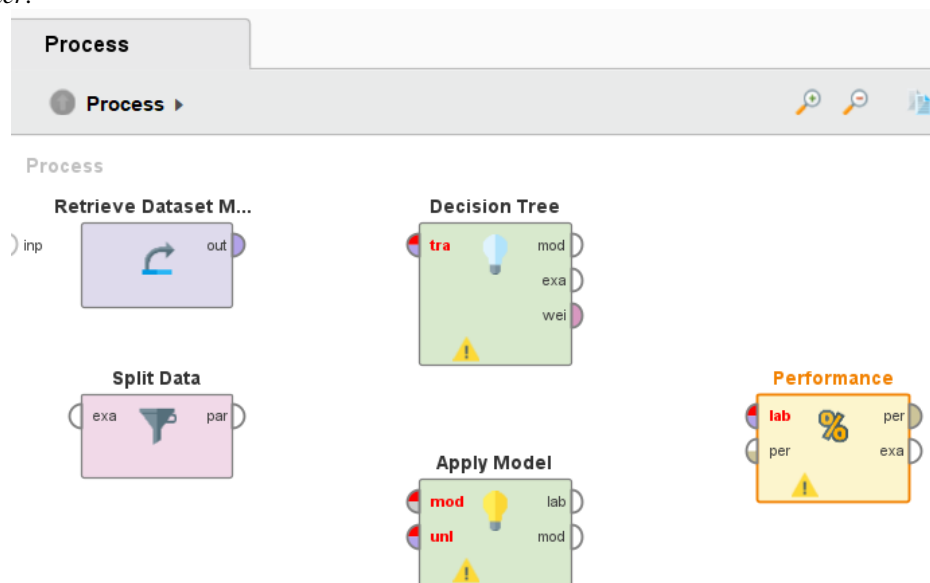
3.6. Proses pembentukan model dalam RapidMiner

Proses selanjutnya adalah *import* data melalui aplikasi *RapidMiner*. Proses ini akan mengatur kondisi atribut sebagai *binominal* dan *polynomial* dilanjutkan dengan pemilihan label yang menjadi hasil keputusan.

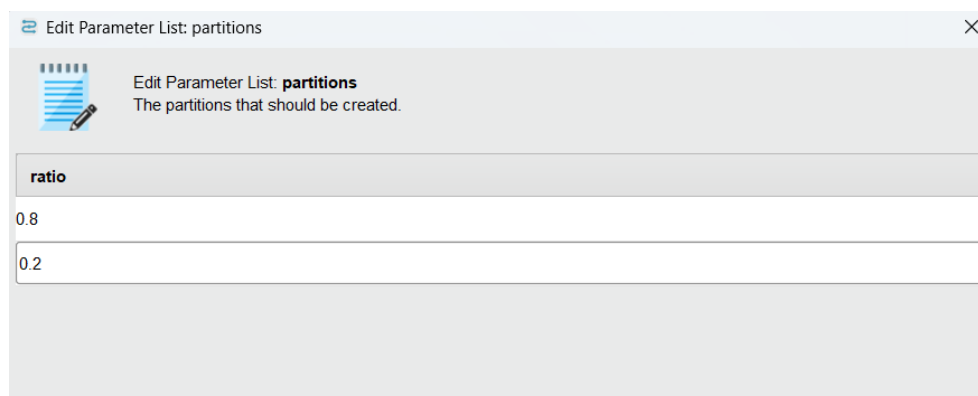
Gambar 5. Import Data

3.7. Persiapan modeling

Dalam lembar kerja *RapidMiner* dipersiapkan dataset, operator *SplitData* (untuk membagi *data training* 80% dan *data testing* 20%). Cukup dilakukan *drag and drop* dan pengaturan dapat dilakukan dalam setiap operator dalam *RapidMiner*.



Gambar 6. Persiapan Modeling



Gambar 7. Split Data Training dan Data Testing

3.8. Modeling

Proses pembentukan model menggunakan parameter kedalaman analisis (*maximal depth*) diujicoba mulai dari kedalaman dua sampai menemukan hasil yang akurasi yang paling optimal. dalam *apply pruning - confidence* diisi 0.5 dan dalam *apply prepruning - minimal leaf size* diisi 3. Pruning adalah menyederhanakan atau memangkas struktur pohon yang telah terbentuk agar percabangan pohon tidak terlalu banyak dan lebih fokus [26].

3.9. Evaluasi Model

Evaluasi dilakukan dengan memberikan kesimpulan atas hasil akurasi yang didapatkan untuk menemukan performa terbaik dan paling efisien. Berdasarkan hasil akurasi, dapat dilihat kesimpulan dari persentase akurasi sebagai berikut [27]:

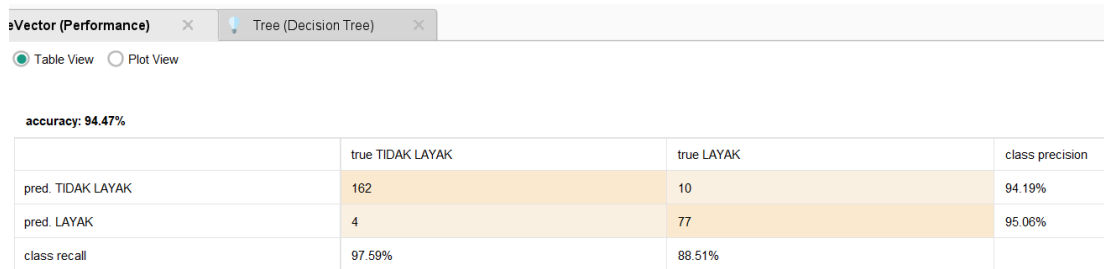
Tabel 2. Standarisasi Akurasi

No.	Hasil Akurasi	Kesimpulan
1.	0.90 – 1.00	Sangat baik sekali
2.	0.80 – 0.90	Sangat baik
3.	0.70 – 0.80	Baik

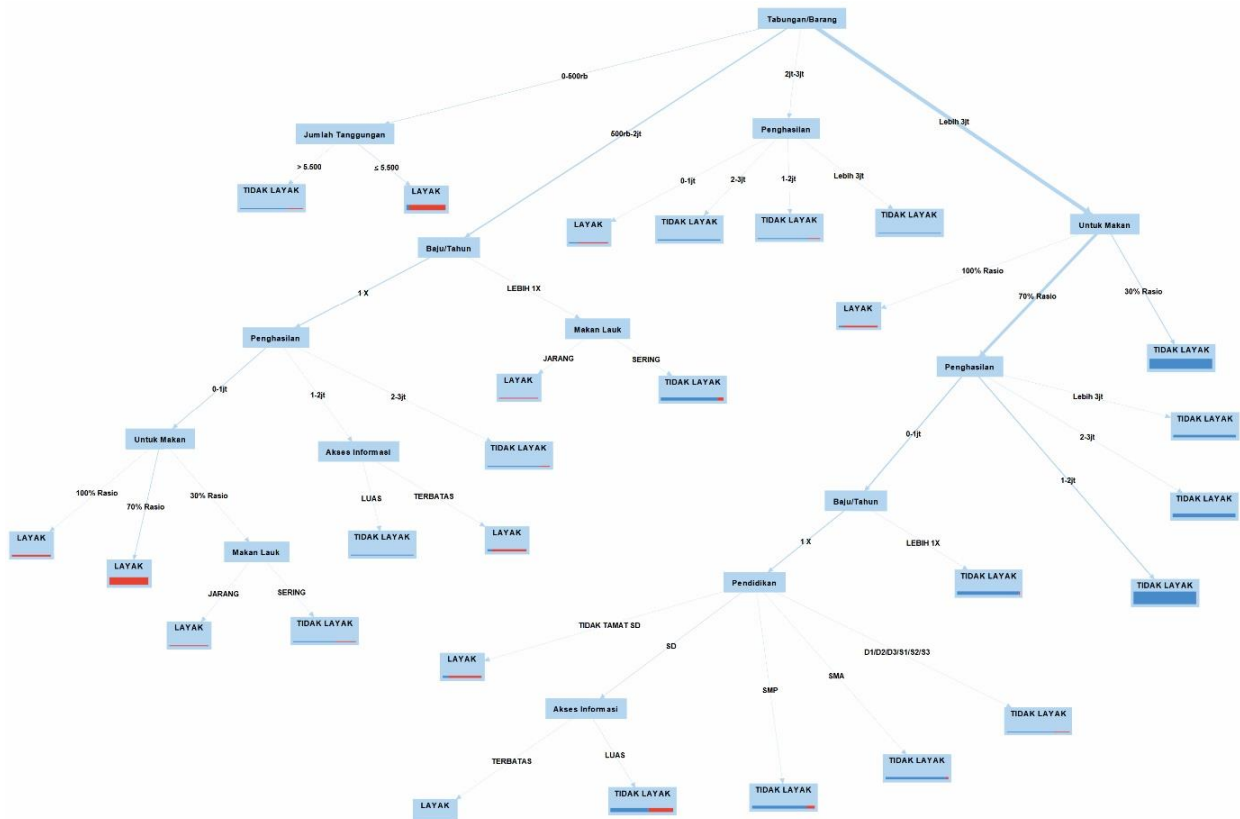
No.	Hasil Akurasi	Kesimpulan
4.	0.60 – 0.70	Cukup
5.	0.50 – 0.60	Sedang

Tabel 3. Pengujian Akurasi Kedalaman Analisis

No.	Kedalaman Cabang	Persentase akurasi	Kesimpulan
1.	2	85,38 %	Sangat baik
2.	3	85,77 %	Sangat baik
3.	4	85,38 %	Sangat baik
4.	5	86,17 %	Sangat baik
5.	6	91,30 %	Sangat baik sekali
6.	7	94,47 %	Sangat baik sekali
7.	8	93,28 %	Sangat baik sekali
8.	9	94,07 %	Sangat baik sekali
9.	10	94,07 %	Sangat baik sekali
10.	11	94,07 %	Sangat baik sekali
11.	12	94,47 %	Sangat baik sekali



Gambar 8. Perma Algoritma C4.5



Gambar 9. Pohon keputusan dengan performa terbaik

3.10. Perbandingan akurasi dengan Algoritma Naïve Bayes

Table View Plot View

accuracy: 93.28%

	true TIDAK LAYAK	true LAYAK	class precision
pred. TIDAK LAYAK	156	7	95.71%
pred. LAYAK	10	80	88.89%
class recall	93.98%	91.95%	

Gambar 10. Performa Algoritma Naïve Bayes

3.11. Perbandingan akurasi dengan Algoritma K-Nearest Neighbors (K-NN)

Table View Plot View

accuracy: 91.70%

	true TIDAK LAYAK	true LAYAK	class precision
pred. TIDAK LAYAK	162	17	90.50%
pred. LAYAK	4	70	94.59%
class recall	97.59%	80.46%	

Gambar 11. Performa Algoritma K-NN

4. SIMPULAN

Berdasarkan hasil analisa sampel Data MCD Desa Sumbung, Kecamatan Cepogo, Kabupaten Boyolali, diperoleh hasil penelitian sebagai berikut:

1. Kondisi ekonomi warga Kabupaten Boyolali paling dipengaruhi oleh: kepemilikan harta lancar, jumlah tanggungan dalam satu keluarga, penghasilan bulanan, kemampuan membeli lauk, rasio pengeluaran pangan, kemampuan membeli pakaian, kepemilikan akses informasi, dan tingkat pendidikan.
2. Pengujian menggunakan *RapidMiner* menghasilkan akurasi sebesar 94,47% untuk C4.5, 93,28% untuk *Naïve Bayes* dan 91,70% untuk *K-Nearest Neighbor (KNN)* sehingga C4.5 lebih cocok untuk diterapkan dalam klasifikasi kondisi ekonomi warga Kabupaten Boyolali berdasarkan data MCD (*Monitoring Center for Development*).

5. SARAN

Dalam penelitian ini diteliti sampel data dari satu desa di Kabupaten Boyolali dan berfokus pada indikator kondisi ekonomi. Penelitian selanjutnya dapat menganalisis tentang kelayakan rumah yang ditinggali warga Boyolali. Data MCD Boyolali diperbarui setiap bulan sehingga dapat dianalisa pola kondisi ekonomi yang akan menurun maupun meningkat. Dalam pengembangan aplikasi berdasarkan penelitian ini pun, sebaiknya benar-benar memperhatikan regulasi yang berlaku dalam instansi.

DAFTAR PUSTAKA

[1] Y. Puspitasari Gobel, “Pemulihan Ekonomi Indonesia Pasca Pandemi Covid-19 Dengan Mengkombinasikan Model Filantropi Islam Dan Ndeas Model,” *J. Tabarru’ Islam. Bank. Financ.*, vol. 3, no. 2, hal. 209–223, 2020, doi: 10.25299/jtb.2020.vol3(2).5809.

[2] W. Lidysari, H. S. Tambunan, dan H. Qurniawan, “Penerapan Data Mining Dalam Menentukan Kelayakan Penerima Bantuan Sosial Pemko Dengan Algoritma C4.5 (Kasus Kantor Kelurahan Martoba),” *Kesatria J. Penerapan Sist. Inf. (Komputer dan Manajemen)*, vol. 3, no. 1, hal. 53–61, 2022, doi:

- 10.30645/kesatria.v3i1.97.
- [3] I. P. Ninditama, W. Cholil, M. Akbar, dan D. Antoni, "Klasifikasi Keluarga Sejahtera Study Kasus: Kecamatan Kota Palembang," *J. TEKNO KOMPAK*, vol. 15, no. 2, hal. 37–49, 2020, doi: <https://doi.org/10.33365/jtk.v15i2.1156>.
 - [4] F. Indahsari, D. Susanti, dan A. Supriyanto, "Implementasi Metode SAW-AHP Dalam Penentuan Prioritas Penerima Bantuan Sosial Kemiskinan Berdasarkan Kriteria BPS," *Jutisi J. Ilm. Tek. Inform. dan Sist. Inf.*, vol. 13, no. 1, hal. 234–246, 2024, doi: <http://dx.doi.org/10.35889/jutisi.v13i1.1766>.
 - [5] S. Wahyuni, E. Hariyanto, dan S. Batubara, *Data Mining dengan Decision Tree C4.5 dan Apriori (Konsep dan Implementasi Menggunakan RapidMiner dan Weka)*, Cetakan 1. Yogyakarta: Penerbit Deepublish (Group Penerbitan CV Budi Utama), 2022.
 - [6] R. Maulana, A. J. Pitoyo, dan M. A. F. Alfana, "Analisis Pengaruh Kemiskinan dan Kondisi Ekonomi Terhadap Indeks Pembangunan Manusia di Provinsi Jawa Tengah Tahun 2013-2017," *Media Komun. Geogr.*, vol. 23, no. 1, hal. 12–24, 2022, doi: 10.23887/mkg.v23i1.39301.
 - [7] Menteri Sosial Republik Indonesia, *Keputusan Menteri Sosial Republik Indonesia Nomor 262/HUK/2022 Tentang Kriteria Fakir Miskin*, no. 31.12.2022. 2022, hal. 1–3. [Daring]. Tersedia pada: <https://dinsos.jogjaprov.go.id/wp-content/uploads/2023/03/KEPMEN-262-TAHUN-2022-Kriteria-Fakir-Miskin.pdf>
 - [8] A. Sitepu, "KARAKTERISTIK KELUARGA MENURUT PERINGKAT KEMISKINAN: STUDI PENDAHULUAN UNTUK PERUMUSAN KRITERIA FAKIR MISKIN," *Pus. Penelit. dan Pengemb. Kesejaht. Sos. Kementeri. Sos. Republik Indones.*, vol. 17, no. 01, hal. 49–63, 2012, doi: <https://ejournal.kemensos.go.id/index.php/Sosioinforma/article/view/930/490>.
 - [9] M. F. Rizqullah, N. T. Raihana, dan M. I. Jambak, "Komparasi Penerapan Algoritma C4.5, K-Nearest Neighbor, dan Naïve Bayes untuk Keberlangsungan Pasien Gagal Jantung," *KLIK Kaji. Ilm. Inform. dan Komput.*, vol. 4, no. 5, hal. 2580–2587, 2024, doi: 10.30865/klik.v4i5.1788.
 - [10] E. Fitriani, "Perbandingan Algoritma C4.5 Dan Naïve Bayes Untuk Menentukan Kelayakan Penerima Bantuan Program Keluarga Harapan," *Sistemasi*, vol. 9, no. 1, hal. 103, 2020, doi: 10.32520/stmsi.v9i1.596.
 - [11] A. Supriyadi, "Perbandingan Algoritma Naive Bayes dan Decision Tree(C4.5) dalam Klasifikasi Dosen Berprestasi," *Gener. J.*, vol. 7, no. 1, hal. 39–49, 2023, doi: 10.29407/gj.v7i1.19797.
 - [12] M. Martin dan L. Nilawati, "Comparison of C4.5 and Naïve Bayes Algorithms for Assessment of Public Complaints Services," *J. Informatics Telecommun. Eng.*, vol. 5, no. 1, hal. 101–111, 2021, doi: 10.31289/jite.v5i1.5292.
 - [13] T. Wijayanti, "CATATAN BERITA - UJDIH BPK Perwakilan Provinsi Jawa Tengah," hal. 1–5, 2023. [Daring]. Tersedia pada: <https://jateng.bpk.go.id/wp-content/uploads/2023/11/73.pdf>
 - [14] S. Pratama, I. Iswandi, A. Sevitan, dan T. P. Anjani, "Penerapan Data Mining Untuk Memprediksi Prestasi Akademik Mahasiswa Menggunakan Algoritma C4.5 dengan CRISP-DM," *J. Appl. Informatics Comput.*, vol. 7, no. 1, hal. 20–14, 2023, doi: 10.30871/jaic.v7i1.4998.
 - [15] S. D. Damanik dan M. I. Jambak, "Klasifikasi Customer Churn pada Telekomunikasi Industri Untuk Retensi Pelanggan Menggunakan Algoritma C4.5," *KLIK Kaji. Ilm. Inform. dan Komput.*, vol. 3, no. 6, hal. 1303–1309, 2023, doi: 10.30865/klik.v3i6.829.
 - [16] G. Fajriansyah, "Analisis Daftar Pemilih Tetap Pada Hasil Rekapitulasi Kpu Berdasarkan Usia Menggunakan Algoritma K-Means (Studi Kasus : Kota Bandar Lampung)," *Electrician*, vol. 15, no. 1, hal. 39–53, 2021, doi: 10.23960/elc.v15n1.2147.
 - [17] N. H. Purnomo, B. Pamungkas, dan C. Juliane, "Penerapan Algoritma C4.5 Untuk Klasifikasi Tren Pelanggaran Kendaraan Angkutan Barang dengan Metode CRISP-DM," *J. Media Inform. Budidarma*, vol. 7, no. 1, hal. 30–40, 2023, doi: 10.30865/mib.v7i1.5247.
 - [18] Y. A. Singgalen, "Analisis Sentimen dan Sistem Pendukung Keputusan Menginap di Hotel Menggunakan Metode CRISP-DM dan SAW," *J. Inf. Syst. Res.*, vol. 4, no. 4, hal. 1343–1353, 2023, doi: 10.47065/josh.v4i4.3917.
 - [19] N. Ratwastuti, L. Atikah, R. Rascaliala, dan Suhendra, "Prediksi Prestasi Belajar Mahasiswa Vokasi Menggunakan Algoritma C4.5," *J. Ilm. Komputasi*, vol. 22, no. 1, hal. 13–19, 2023, doi: 10.32409/jikstik.22.1.3320.
 - [20] J. Santos-Pereira, L. Gruenwald, dan J. Bernardino, "Top data mining tools for the healthcare industry," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 34, no. 8, hal. 4968–4982, 2022, doi: 10.1016/j.jksuci.2021.06.002.
 - [21] H. Hendri, "Implementasi Data Mining Dengan Metode C4.5 Untuk Prediksi Mahasiswa Penerima Beasiswa," *Indones. J. Comput. Sci.*, vol. 10, no. 2, hal. 312–321, 2021, doi: 10.33022/ijcs.v10i2.3013.
 - [22] S. D. Manullang, E. Buulolo, dan I. Lubis, "Implementasi Data Mining Dalam Memprediksi Jumlah Pinjaman Dengan Algoritma C4.5 Pada Kopdit CU Damai Sejahtera," *J. Sist. Komput. dan Inform.*, vol. 1, no. 3, hal. 265, 2020, doi: 10.30865/json.v1i3.2153.
 - [23] D. Puspita, S. Aminah, dan A. Arif, "Prediction System for Credit Eligibility Using C4.5 Algorithm," *J.*

- Informatics Telecommun. Eng.*, vol. 6, no. 1, hal. 148–156, 2022, doi: 10.31289/jite.v6i1.7311.
- [24] N. Leniawati dan S. Wijayanto, “Klasifikasi Desa Wisata di Provinsi Jawa Tengah Menggunakan Algoritma C4.5,” *J. Sains Komput. Inform.*, vol. 8, no. 1, hal. 171–184, 2024, doi: <http://dx.doi.org/10.30645/j-sakti.v8i1>.
- [25] B. R. Heryanti, “Implementasi Perubahan Kebijakan Batas Usia Perkawinan,” *J. Ius Const.*, vol. 6, no. 1, hal. 120, 2021, doi: 10.26623/jic.v6i1.3190.
- [26] Y. Bastian, H. Satria Tambunan, dan W. Saputra, “Analisis Penerapan Algoritma C4.5 Dalam Mengukur Tingkat Kepuasan Pelanggan Indihome Pada Kota Pematangsiantar,” *KESATRIA J. Penerapan Sist. Inf. (Komputer Manajemen)*, vol. 2, no. 1, hal. 62–69, 2021, doi: <https://doi.org/10.30645/kesatria.v2i1.59>.
- [27] S. Dewi dan O. Oktaviawati, “Penerapan Algoritma C4.5 untuk Pehamanan Siswa SMK Pada Pelajaran Kompetensi Keahlian,” *Intern. Inf. Syst. J.*, vol. 5, no. 2, hal. 116–125, 2022, doi: 10.32627.