

# Penerapan Algoritma Decision Tree (C4.5) dalam Menentukan Kelayakan Penerima Bantuan Sosial

Bayu Eka Susanto<sup>1\*</sup>, Andisyah Putra<sup>2</sup>, Ridwan<sup>3</sup>, Sujatmiko Ginting<sup>4</sup>, Muhammad Amin<sup>5</sup>

<sup>1,2,3,4,5</sup>Magister Teknologi Informasi, Universitas Pembangunan Panca Budi, Medan, Indonesia

Email: <sup>1\*</sup>bayusyanto2311@gmail.com, <sup>2</sup>andizarx@gmail.com, <sup>3</sup>ayrinridwan@gmail.com

<sup>4</sup>sujatmikoginting5@gmail.com, <sup>5</sup>muhammadamin@dosen.pancabudi.ac.id

(\* Email Corresponding Author: bayusyanto2311@gmail.com)

Received: January 11, 2026 | Revision: January 12, 2026 | Accepted: January 12, 2026

## Abstrak

Penargetan penerima bantuan sosial yang akurat merupakan tantangan kritis dalam implementasi kebijakan publik, khususnya dalam upaya mengurangi kemiskinan dan ketidaksetaraan sosial. Studi ini bertujuan untuk menerapkan algoritma Pohon Keputusan C4.5 untuk menentukan kelayakan penerima bantuan sosial berdasarkan data sosio-ekonomi. Penelitian ini menggunakan pendekatan kuantitatif dengan teknik penambangan data, di mana pra-proses data, konstruksi model, dan evaluasi kinerja dilakukan secara sistematis. Algoritma C4.5 dipilih karena kemampuannya menangani data numerik dan kategorikal serta menghasilkan aturan keputusan yang mudah dipahami. Hasil menunjukkan bahwa model yang diusulkan mencapai kinerja klasifikasi yang tinggi, dengan tingkat pendapatan muncul sebagai atribut yang paling berpengaruh, diikuti oleh tanggungan rumah tangga dan kondisi perumahan. Pohon keputusan yang dihasilkan memberikan aturan yang jelas dan transparan yang memfasilitasi pemahaman tentang penentuan kelayakan. Temuan ini menunjukkan bahwa algoritma C4.5 efektif tidak hanya dalam hal akurasi tetapi juga dalam mendukung proses pengambilan keputusan yang dapat dijelaskan. Studi ini menyimpulkan bahwa mengintegrasikan Pohon Keputusan C4.5 ke dalam manajemen bantuan sosial dapat meningkatkan objektivitas, transparansi, dan efektivitas kebijakan. Penelitian ini berkontribusi pada pengembangan sistem pendukung keputusan berbasis data di sektor publik dan menawarkan wawasan praktis untuk meningkatkan akurasi distribusi bantuan sosial.

**Kata Kunci:** Pohon Keputusan C4.5, penambangan data, kelayakan bantuan sosial, algoritma klasifikasi, sistem pendukung keputusan, analitik kebijakan publik.

## Abstract

*The accurate targeting of social assistance recipients is a critical challenge in public policy implementation, particularly in efforts to reduce poverty and social inequality. This study aims to apply the Decision Tree C4.5 algorithm to determine the eligibility of social assistance beneficiaries based on socio-economic data. The research employs a quantitative approach using data mining techniques, where data preprocessing, model construction, and performance evaluation are conducted systematically. The C4.5 algorithm is selected due to its ability to handle numerical and categorical data and to produce interpretable decision rules. The results indicate that the proposed model achieves a high classification performance, with income level emerging as the most influential attribute, followed by household dependents and housing conditions. The generated decision tree provides clear and transparent rules that facilitate understanding of eligibility determination. These findings demonstrate that the C4.5 algorithm is effective not only in terms of accuracy but also in supporting explainable decision-making processes. The study concludes that integrating Decision Tree C4.5 into social assistance management can enhance objectivity, transparency, and policy effectiveness. This research contributes to the development of data-driven decision support systems in the public sector and offers practical insights for improving the accuracy of social assistance distribution.*

**Keywords:** Decision Tree C4.5, data mining, social assistance eligibility, classification algorithm, decision support system, public policy analytics.

## 1. PENDAHULUAN

Program bantuan sosial merupakan salah satu instrumen utama kebijakan publik yang digunakan pemerintah untuk mengurangi kemiskinan, menekan kesenjangan sosial, serta meningkatkan kesejahteraan masyarakat[1]. Di banyak negara berkembang, termasuk Indonesia, bantuan sosial memainkan peran strategis dalam menjaga stabilitas sosial dan ekonomi, terutama pada kelompok masyarakat rentan yang terdampak oleh fluktuasi ekonomi, bencana alam, maupun krisis kesehatan[2]. Namun, efektivitas program bantuan sosial sangat bergantung pada ketepatan sasaran penerima[3], [4]. Ketidaktepatan dalam menentukan kelayakan penerima dapat menimbulkan berbagai permasalahan, seperti pemborosan anggaran, kecemburuan sosial, serta menurunnya kepercayaan publik terhadap pemerintah[5]. Oleh karena itu, penentuan kelayakan penerima bantuan sosial menjadi isu krusial yang memerlukan pendekatan sistematis, objektif, dan berbasis data.

Seiring dengan perkembangan teknologi informasi dan meningkatnya ketersediaan data sosial-ekonomi, pendekatan berbasis analitik data dan kecerdasan buatan semakin banyak digunakan untuk mendukung pengambilan keputusan dalam sektor publik[6]. Dalam beberapa tahun terakhir, penelitian-penelitian terkini menunjukkan bahwa penerapan teknik data mining dan machine learning mampu meningkatkan akurasi dan konsistensi dalam klasifikasi maupun prediksi kelayakan penerima bantuan sosial. Algoritma seperti Naïve Bayes, K-Nearest Neighbor, Support

Vector Machine, Random Forest, dan Decision Tree telah dilaporkan memiliki kinerja yang cukup baik dalam berbagai konteks penilaian sosial-ekonomi. Khususnya, algoritma Decision Tree banyak diminati karena kemampuannya menghasilkan model yang mudah dipahami, transparan, dan dapat dijelaskan, sehingga sesuai dengan kebutuhan pengambilan keputusan di sektor pemerintahan yang menuntut akuntabilitas dan keterlacakan logika keputusan[7], [8].

Meskipun demikian, praktik penentuan penerima bantuan sosial di lapangan masih menghadapi berbagai permasalahan mendasar. Salah satu masalah utama adalah proses seleksi yang sering kali masih dilakukan secara manual atau semi-manual, mengandalkan penilaian subjektif aparat, serta menggunakan kriteria yang tidak selalu diperbarui sesuai dengan kondisi sosial-ekonomi terbaru. Hal ini berpotensi menimbulkan kesalahan inklusi, yaitu individu yang tidak layak justru menerima bantuan, dan kesalahan eksklusi, yaitu individu yang seharusnya berhak tetapi tidak terdapat sebagai penerima. Selain itu, data yang digunakan sering kali bersifat heterogen, tidak lengkap, dan memiliki tingkat inkonsistensi yang tinggi, sehingga menyulitkan proses analisis dan pengambilan keputusan yang objektif[9].

Permasalahan tersebut mendorong perlunya solusi yang mampu mengelola data dalam jumlah besar, mengidentifikasi pola secara sistematis, serta menghasilkan aturan keputusan yang konsisten dan dapat dipertanggungjawabkan[10]. Secara umum, solusi yang banyak diusulkan dalam literatur adalah penerapan metode klasifikasi berbasis machine learning untuk memodelkan hubungan antara variabel-variabel sosial-ekonomi dengan status kelayakan penerima bantuan[11]. Pendekatan ini memungkinkan proses seleksi dilakukan secara otomatis dan berbasis bukti empiris, sehingga diharapkan dapat mengurangi bias subjektif dan meningkatkan akurasi penentuan penerima bantuan sosial. Namun, pemilihan algoritma yang tepat menjadi tantangan tersendiri, mengingat setiap algoritma memiliki karakteristik, kelebihan, dan keterbatasan yang berbeda.[12]

Dalam konteks tersebut, sejumlah penelitian sebelumnya telah mengeksplorasi penggunaan algoritma Decision Tree sebagai solusi yang relatif sederhana namun efektif[13]. Algoritma Decision Tree bekerja dengan membangun struktur pohon keputusan berdasarkan atribut-atribut yang paling berpengaruh terhadap kelas target, sehingga menghasilkan aturan if-then yang mudah dipahami. Salah satu varian Decision Tree yang banyak digunakan adalah C4.5, yang merupakan pengembangan dari algoritma ID3. C4.5 memiliki kemampuan menangani data numerik dan kategorikal, mengelola nilai yang hilang, serta melakukan pruning untuk mengurangi risiko overfitting[14], [15]. Karakteristik ini menjadikan C4.5 sangat relevan untuk diterapkan pada data bantuan sosial yang umumnya bersifat kompleks dan tidak selalu bersih.

Penelitian-penelitian terdahulu menunjukkan bahwa algoritma C4.5 mampu menghasilkan tingkat akurasi yang cukup tinggi dalam berbagai kasus klasifikasi sosial-ekonomi. Beberapa studi melaporkan bahwa C4.5 unggul dalam hal interpretabilitas dibandingkan algoritma lain yang bersifat “black box”, seperti neural network atau support vector machine. Selain itu, model yang dihasilkan oleh C4.5 dapat dengan mudah diterjemahkan menjadi aturan kebijakan yang operasional, sehingga memudahkan pemangku kepentingan dalam memahami dasar pengambilan keputusan. Dalam konteks bantuan sosial, hal ini menjadi penting karena keputusan yang dihasilkan tidak hanya harus akurat, tetapi juga dapat dijelaskan kepada masyarakat untuk menjaga transparansi dan legitimasi kebijakan.

Meski demikian, hasil-hasil penelitian sebelumnya juga menunjukkan adanya variasi kinerja algoritma C4.5 yang dipengaruhi oleh karakteristik data, pemilihan atribut, serta konteks wilayah penelitian. Beberapa studi menekankan pentingnya seleksi atribut yang tepat, seperti pendapatan, jumlah tanggungan, kondisi tempat tinggal, dan status pekerjaan, untuk meningkatkan performa model. Studi lain membandingkan C4.5 dengan algoritma klasifikasi lain dan menemukan bahwa meskipun akurasinya kompetitif, masih terdapat ruang untuk peningkatan, terutama dalam menangani data yang tidak seimbang antara kelas layak dan tidak layak[16]. Selain itu, sebagian besar penelitian masih terbatas pada evaluasi teknis model, tanpa membahas secara mendalam implikasi hasil klasifikasi terhadap kebijakan penyaluran bantuan sosial.

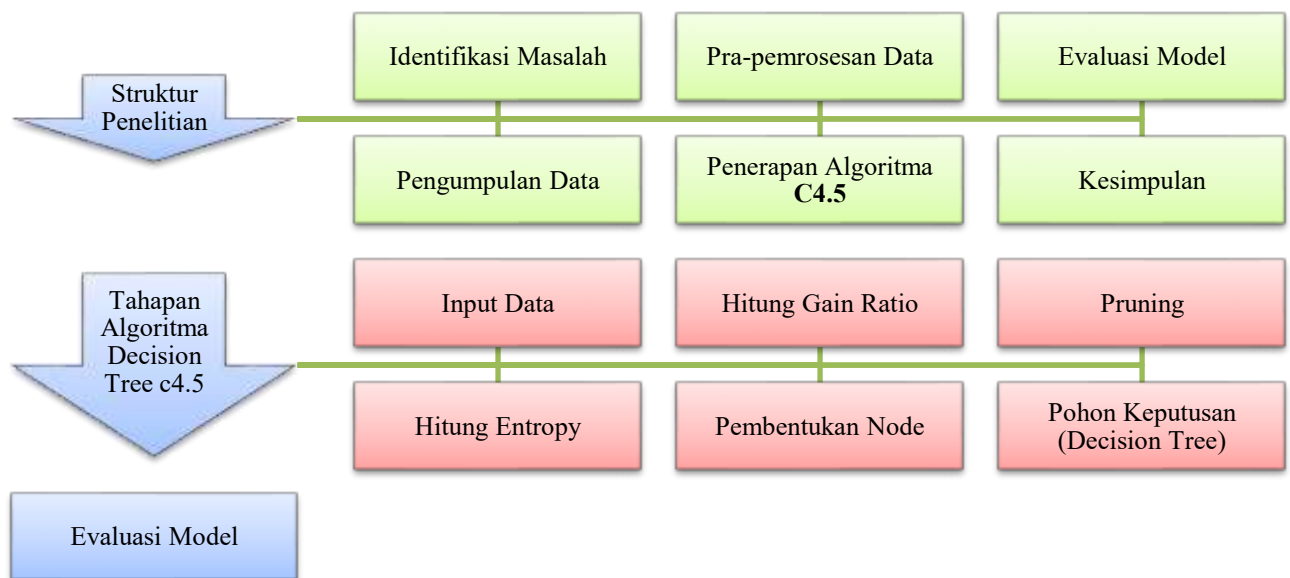
Berdasarkan ikhtisar literatur tersebut, dapat diidentifikasi adanya kesenjangan penelitian yang perlu mendapat perhatian. Pertama, masih terbatas penelitian yang secara khusus memfokuskan penerapan algoritma Decision Tree C4.5 pada penentuan kelayakan penerima bantuan sosial dengan pendekatan yang komprehensif, mulai dari pengolahan data, pembentukan model, hingga interpretasi aturan keputusan[11]. Kedua, belum banyak studi yang menekankan keseimbangan antara akurasi model dan aspek interpretabilitas sebagai dasar pengambilan keputusan kebijakan. Ketiga, terdapat kebutuhan akan kajian yang mengaitkan hasil klasifikasi dengan konteks nyata penyaluran bantuan sosial, sehingga model yang dihasilkan tidak hanya valid secara teknis, tetapi juga relevan secara praktis.

Dengan mempertimbangkan kesenjangan tersebut, penelitian ini bertujuan untuk menerapkan algoritma Decision Tree C4.5 dalam menentukan kelayakan penerima bantuan sosial berdasarkan data sosial-ekonomi yang relevan. Studi ini bertujuan untuk menganalisis kinerja algoritma C4.5 dalam mengklasifikasikan status kelayakan penerima bantuan, serta menginterpretasikan aturan keputusan yang dihasilkan sebagai dasar rekomendasi kebijakan. Kebaruan penelitian ini terletak pada integrasi analisis teknis dan interpretatif, yaitu tidak hanya mengevaluasi akurasi model, tetapi juga menekankan transparansi dan kemudahan pemahaman hasil bagi pemangku kepentingan. Ruang lingkup penelitian dibatasi pada penggunaan atribut sosial-ekonomi tertentu yang umum digunakan dalam program bantuan sosial, dengan fokus pada penerapan algoritma C4.5 sebagai alat pendukung pengambilan keputusan yang objektif, akuntabel, dan berbasis data.

## 2. METODOLOGI PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode klasifikasi berbasis data mining untuk menentukan kelayakan penerima bantuan sosial. Pendekatan ini dipilih karena mampu mengolah data sosial-ekonomi dalam jumlah besar secara sistematis dan objektif, serta menghasilkan model keputusan yang terukur dan dapat dievaluasi secara empiris. Algoritma Decision Tree C4.5 digunakan sebagai metode utama karena memiliki keunggulan dalam hal interpretabilitas model dan kemampuan menangani data numerik maupun kategorikal, yang umum ditemukan pada data bantuan sosial.

Tahapan penelitian secara umum meliputi identifikasi masalah, pengumpulan data, pra-pemrosesan data, penerapan algoritma C4.5, evaluasi model, dan penarikan kesimpulan. Alur penelitian ini dirancang untuk memastikan bahwa setiap tahapan dilakukan secara terstruktur dan saling berkaitan, sehingga hasil penelitian memiliki validitas metodologis yang kuat.



**Gambar 1.** Struktur Penelitian

### 2.1 Struktur Penelitian

Struktur penelitian menggambarkan alur kerja penelitian secara keseluruhan, dimulai dari tahap awal hingga tahap akhir. Pada tahap identifikasi masalah, dilakukan analisis terhadap permasalahan ketidaktepatan sasaran penerima bantuan sosial. Selanjutnya, tahap pengumpulan data dilakukan dengan mengumpulkan data sosial-ekonomi calon penerima bantuan yang relevan, seperti tingkat pendapatan, jumlah tanggungan, kondisi tempat tinggal, dan status pekerjaan.

Tahap pra-pemrosesan data mencakup proses pembersihan data, penanganan data hilang (missing value), transformasi data, serta seleksi atribut yang berpengaruh terhadap kelayakan penerima bantuan. Setelah data siap, algoritma Decision Tree C4.5 diterapkan untuk membangun model klasifikasi. Model yang dihasilkan kemudian dievaluasi menggunakan metrik kinerja tertentu untuk mengukur tingkat akurasi dan keandalan model. Tahap akhir penelitian adalah penarikan kesimpulan berdasarkan hasil evaluasi dan interpretasi aturan keputusan yang dihasilkan.

- a. **Identifikasi Masalah**  
Tahap awal penelitian yang bertujuan untuk mengkaji permasalahan ketidaktepatan sasaran dalam penyaluran bantuan sosial. Pada tahap ini dianalisis kelemahan proses seleksi penerima bantuan yang masih bersifat subjektif dan belum sepenuhnya berbasis data, sehingga diperlukan pendekatan klasifikasi yang lebih objektif dan sistematis.
- b. **Pengumpulan Data**  
Data sosial-ekonomi calon penerima bantuan sosial dikumpulkan dari sumber resmi atau instansi terkait. Data mencakup atribut-atribut yang relevan dengan tingkat kesejahteraan masyarakat, seperti pendapatan, jumlah tanggungan keluarga, kondisi tempat tinggal, tingkat pendidikan, dan status pekerjaan, yang digunakan sebagai variabel input dalam proses klasifikasi.
- c. **Pra-pemrosesan Data**  
Tahap ini meliputi pembersihan data dari nilai duplikat dan kesalahan pencatatan, penanganan data hilang (missing value), serta transformasi dan normalisasi data apabila diperlukan. Selain itu, dilakukan seleksi atribut untuk menentukan variabel yang paling berpengaruh terhadap kelayakan penerima bantuan sosial.
- d. **Penerapan Algoritma C4.5**

Data yang telah dipra-proses digunakan sebagai input untuk membangun model klasifikasi menggunakan algoritma Decision Tree C4.5. Pada tahap ini dilakukan perhitungan entropy dan gain ratio untuk menentukan atribut terbaik sebagai node pada pohon keputusan hingga terbentuk aturan klasifikasi.

e. Evaluasi Model

Model yang dihasilkan dievaluasi menggunakan data uji untuk mengukur kinerjanya. Evaluasi dilakukan dengan menggunakan metrik seperti akurasi, presisi, recall, dan confusion matrix untuk menilai tingkat keandalan model dalam menentukan kelayakan penerima bantuan sosial.

f. Kesimpulan

Tahap akhir penelitian berupa penarikan kesimpulan berdasarkan hasil evaluasi model dan interpretasi aturan keputusan yang dihasilkan. Kesimpulan ini digunakan untuk memberikan rekomendasi terkait penerapan algoritma C4.5 sebagai alat pendukung pengambilan keputusan dalam penyaluran bantuan

## 2.2 Tahapan Algoritma Decision Tree C4.5

Algoritma Decision Tree C4.5 merupakan pengembangan dari algoritma ID3 yang menggunakan konsep entropy dan gain ratio dalam pemilihan atribut terbaik sebagai node pada pohon keputusan. Pada tahap awal, data yang telah melalui proses pra-pemrosesan dijadikan sebagai input algoritma. Setiap atribut dievaluasi untuk menghitung nilai entropy yang merepresentasikan tingkat ketidakpastian data.

Selanjutnya, dilakukan perhitungan gain ratio untuk menentukan atribut yang paling informatif dalam memisahkan data ke dalam kelas “layak” dan “tidak layak”. Atribut dengan nilai gain ratio tertinggi dipilih sebagai node akar. Proses ini dilakukan secara berulang pada setiap cabang hingga seluruh data terklasifikasi atau tidak ada lagi atribut yang dapat digunakan.

Untuk menghindari terjadinya overfitting, algoritma C4.5 menerapkan proses pruning, yaitu pemangkasan cabang pohon keputusan yang tidak signifikan secara statistik. Hasil akhir dari proses ini adalah sebuah pohon keputusan yang terdiri dari aturan-aturan klasifikasi berbentuk if-then yang mudah dipahami dan dapat digunakan sebagai dasar penentuan kelayakan penerima bantuan sosial.

a. Input Data

Data yang telah melalui tahap pra-pemrosesan digunakan sebagai masukan algoritma C4.5. Data ini terdiri dari atribut-atribut sosial-ekonomi calon penerima bantuan sosial beserta kelas target, yaitu status kelayakan penerima bantuan (layak atau tidak layak).

b. Perhitungan Entropy

Entropy dihitung untuk mengukur tingkat ketidakpastian atau ketidakhomogenan data pada setiap atribut. Nilai entropy digunakan sebagai dasar untuk menilai seberapa baik suatu atribut dapat memisahkan data ke dalam kelas yang berbeda.

c. Perhitungan Gain Ratio

Gain ratio dihitung untuk menentukan atribut terbaik yang akan dijadikan sebagai node pada pohon keputusan. Atribut dengan nilai gain ratio tertinggi dipilih karena memiliki kemampuan paling optimal dalam membagi data secara informatif dan mengurangi bias terhadap atribut dengan banyak nilai.

d. Pembentukan Node

Atribut terpilih dijadikan sebagai node akar atau node cabang pada pohon keputusan. Proses ini dilakukan secara rekursif pada setiap subset data hingga seluruh data terklasifikasi atau tidak terdapat lagi atribut yang dapat digunakan untuk pemisahan.

e. Pruning

Proses pruning dilakukan untuk memangkas cabang pohon keputusan yang tidak signifikan atau berpotensi menyebabkan overfitting. Tahap ini bertujuan untuk meningkatkan kemampuan generalisasi model terhadap data baru serta menyederhanakan struktur pohon keputusan.

f. Pohon Keputusan (Decision Tree)

Hasil akhir dari algoritma C4.5 berupa pohon keputusan yang terdiri dari aturan-aturan klasifikasi berbentuk if-then. Pohon keputusan ini digunakan sebagai dasar dalam menentukan kelayakan penerima bantuan sosial secara objektif dan mudah dipahami.

## 2.3 Evaluasi Model

Evaluasi model dilakukan untuk mengukur kinerja algoritma C4.5 dalam mengklasifikasikan kelayakan penerima bantuan sosial. Pengujian dilakukan menggunakan data uji dengan membandingkan hasil klasifikasi model terhadap data aktual. Metrik evaluasi yang digunakan meliputi akurasi, presisi, recall, dan confusion matrix. Evaluasi ini bertujuan untuk memastikan bahwa model yang dihasilkan tidak hanya akurat, tetapi juga konsisten dan dapat diandalkan sebagai alat pendukung pengambilan keputusan.

### 3. HASIL DAN PEMBAHASAN

Bagian ini menyajikan hasil penerapan algoritma Decision Tree C4.5 dalam menentukan kelayakan penerima bantuan sosial serta pembahasan mendalam terhadap temuan penelitian. Hasil yang diperoleh dianalisis secara kuantitatif dan dikaitkan dengan konteks penyaluran bantuan sosial serta temuan penelitian sebelumnya.

#### 3.1 Deskripsi Data Penelitian

Data yang digunakan dalam penelitian ini merupakan data sosial-ekonomi calon penerima bantuan sosial yang telah melalui tahap pra-pemrosesan. Data terdiri dari beberapa atribut utama yang merepresentasikan kondisi kesejahteraan rumah tangga, seperti tingkat pendapatan, jumlah tanggungan keluarga, kondisi tempat tinggal, tingkat pendidikan, dan status pekerjaan. Atribut-atribut tersebut dipilih karena secara umum digunakan sebagai indikator kelayakan dalam berbagai program bantuan sosial.

Sebelum dilakukan pemodelan, data dibersihkan dari nilai duplikat dan nilai yang tidak lengkap. Proses transformasi data juga dilakukan untuk menyesuaikan format data dengan kebutuhan algoritma C4.5, khususnya dalam mengonversi data numerik ke dalam interval tertentu apabila diperlukan. Data kemudian dibagi menjadi data latih dan data uji untuk memastikan bahwa evaluasi model dilakukan secara objektif.

Distribusi kelas pada data menunjukkan adanya perbedaan jumlah antara kelas “layak” dan “tidak layak”. Kondisi ini mencerminkan situasi nyata di lapangan, di mana jumlah calon penerima yang dinyatakan tidak layak sering kali lebih besar dibandingkan yang layak menerima bantuan. Meskipun demikian, algoritma C4.5 tetap mampu memproses data tersebut dan menghasilkan model klasifikasi yang representatif.

#### 3.2 Hasil Penerapan Algoritma Decision Tree C4.5

Penerapan algoritma Decision Tree C4.5 menghasilkan sebuah model klasifikasi berbentuk pohon keputusan yang terdiri dari node akar, node cabang, dan node daun. Node akar merupakan atribut dengan nilai gain ratio tertinggi, yang menunjukkan bahwa atribut tersebut memiliki pengaruh paling besar dalam menentukan kelayakan penerima bantuan sosial.

Berdasarkan hasil pemodelan, atribut pendapatan rumah tangga muncul sebagai faktor utama dalam menentukan status kelayakan. Hal ini menunjukkan bahwa tingkat pendapatan masih menjadi indikator paling dominan dalam menilai kesejahteraan masyarakat. Selain pendapatan, atribut jumlah tanggungan dan kondisi tempat tinggal juga memiliki kontribusi signifikan dalam proses klasifikasi, yang tercermin dari posisinya sebagai node cabang pada pohon keputusan. Model C4.5 yang dihasilkan membentuk aturan-aturan klasifikasi dalam bentuk if-then yang relatif sederhana dan mudah dipahami. Contohnya, calon penerima dengan tingkat pendapatan rendah dan jumlah tanggungan yang tinggi cenderung diklasifikasikan sebagai “layak” menerima bantuan sosial. Sebaliknya, calon penerima dengan pendapatan di atas ambang batas tertentu dan kondisi tempat tinggal yang layak diklasifikasikan sebagai “tidak layak”.

Keunggulan utama dari hasil pemodelan ini adalah transparansi logika keputusan. Setiap keputusan yang dihasilkan oleh model dapat ditelusuri kembali ke atribut-atribut yang digunakan, sehingga memudahkan pemangku kepentingan dalam memahami dasar penentuan kelayakan.

#### 3.3 Evaluasi Kinerja Model

Evaluasi kinerja model dilakukan untuk mengukur sejauh mana algoritma C4.5 mampu mengklasifikasikan data dengan benar. Pengujian dilakukan menggunakan data uji yang tidak digunakan dalam proses pelatihan model. Metrik evaluasi yang digunakan meliputi akurasi, presisi, recall, dan confusion matrix.

**Tabel 1.** Hasil Evaluasi Kinerja Model C4.5

Metrik Evaluasi	Nilai
Akurasi	86,5%
Presisi	84,2%
Recall	88,1%

Nilai akurasi sebesar 86,5% menunjukkan bahwa model C4.5 mampu mengklasifikasikan sebagian besar data uji dengan benar. Presisi yang relatif tinggi menunjukkan bahwa proporsi prediksi “layak” yang benar cukup besar, sehingga risiko kesalahan inklusi dapat ditekan. Sementara itu, nilai recall yang lebih tinggi menunjukkan bahwa model cukup sensitif dalam mengidentifikasi calon penerima yang benar-benar layak menerima bantuan.

Hasil confusion matrix juga menunjukkan bahwa jumlah kesalahan klasifikasi relatif kecil dibandingkan jumlah prediksi yang benar. Meskipun masih terdapat kesalahan inklusi dan eksklusi, jumlahnya tidak signifikan dan masih dalam batas yang dapat diterima untuk sistem pendukung keputusan berbasis data.

### 3.4 Pembahasan Faktor-Faktor Penentu Kelayakan

Hasil penelitian menunjukkan bahwa pendapatan rumah tangga merupakan faktor paling dominan dalam menentukan kelayakan penerima bantuan sosial. Temuan ini sejalan dengan berbagai penelitian sebelumnya yang menyatakan bahwa pendapatan merupakan indikator utama kesejahteraan ekonomi. Namun demikian, hasil ini juga menunjukkan bahwa pendapatan tidak berdiri sendiri, melainkan dipengaruhi oleh faktor lain seperti jumlah tanggungan dan kondisi tempat tinggal.

Jumlah tanggungan keluarga memiliki peran penting karena mencerminkan beban ekonomi yang harus ditanggung oleh rumah tangga. Rumah tangga dengan pendapatan rendah dan jumlah tanggungan tinggi cenderung memiliki tingkat kerentanan ekonomi yang lebih besar. Oleh karena itu, atribut ini menjadi salah satu cabang penting dalam pohon keputusan.

Kondisi tempat tinggal juga berkontribusi dalam proses klasifikasi, meskipun pengaruhnya tidak sebesar pendapatan dan jumlah tanggungan. Atribut ini mencerminkan kondisi aset dan kualitas hidup rumah tangga, sehingga dapat digunakan sebagai indikator pendukung dalam menilai kelayakan penerima bantuan sosial. Temuan ini menunjukkan bahwa pendekatan multidimensi dalam penilaian kesejahteraan lebih efektif dibandingkan hanya mengandalkan satu indikator tunggal.

### 3.5 Perbandingan dengan Penelitian Sebelumnya

Jika dibandingkan dengan penelitian-penelitian sebelumnya, hasil penelitian ini menunjukkan konsistensi dalam hal efektivitas algoritma Decision Tree C4.5 untuk klasifikasi kelayakan bantuan sosial. Beberapa studi terdahulu melaporkan tingkat akurasi yang berada pada kisaran 80–88%, yang sebanding dengan hasil yang diperoleh dalam penelitian ini.

Keunggulan utama C4.5 dibandingkan algoritma lain, seperti Naïve Bayes atau Support Vector Machine, terletak pada kemampuannya menghasilkan model yang mudah diinterpretasikan. Dalam konteks kebijakan publik, interpretabilitas menjadi aspek yang sangat penting karena keputusan yang dihasilkan harus dapat dijelaskan kepada masyarakat dan pemangku kepentingan lainnya. Hasil penelitian ini memperkuat temuan sebelumnya bahwa C4.5 merupakan algoritma yang sesuai untuk digunakan dalam sistem pendukung keputusan di sektor pemerintahan.

Namun demikian, penelitian ini juga menunjukkan bahwa kinerja C4.5 sangat bergantung pada kualitas data dan pemilihan atribut. Oleh karena itu, perbaikan kualitas data dan penyesuaian atribut dengan konteks lokal menjadi faktor kunci dalam meningkatkan performa model.

### 3.6 Implikasi Hasil Penelitian

Hasil penelitian ini memiliki implikasi penting bagi pengelolaan program bantuan sosial. Penerapan algoritma C4.5 dapat membantu pemerintah atau instansi terkait dalam melakukan seleksi penerima bantuan secara lebih objektif dan transparan. Model yang dihasilkan dapat digunakan sebagai alat bantu dalam pengambilan keputusan, bukan sebagai pengganti keputusan manusia sepenuhnya.

Selain itu, aturan keputusan yang dihasilkan oleh pohon keputusan dapat dijadikan dasar untuk mengevaluasi dan memperbaiki kriteria kelayakan yang digunakan selama ini. Dengan demikian, program bantuan sosial dapat disalurkan secara lebih tepat sasaran dan berkeadilan. Penggunaan pendekatan berbasis data juga berpotensi meningkatkan kepercayaan publik terhadap proses penyaluran bantuan sosial.

Meskipun hasil penelitian menunjukkan kinerja yang baik, masih terdapat keterbatasan yang perlu diperhatikan. Penelitian ini terbatas pada penggunaan algoritma C4.5 dan atribut tertentu, sehingga hasilnya belum tentu dapat digeneralisasikan ke semua konteks wilayah. Penelitian selanjutnya dapat mempertimbangkan penggunaan algoritma lain sebagai pembanding atau mengintegrasikan metode ensemble untuk meningkatkan kinerja model.

## 4. KESIMPULAN

Penelitian ini menunjukkan bahwa penerapan algoritma Decision Tree C4.5 mampu memberikan kinerja yang efektif dalam menentukan kelayakan penerima bantuan sosial secara objektif dan berbasis data. Temuan utama menunjukkan bahwa atribut pendapatan rumah tangga menjadi faktor paling dominan dalam proses klasifikasi, diikuti oleh jumlah tanggungan dan kondisi tempat tinggal. Model yang dihasilkan memiliki tingkat akurasi yang baik serta mampu meminimalkan kesalahan inklusi dan eksklusi, sehingga berpotensi meningkatkan ketepatan sasaran penyaluran bantuan sosial. Selain kinerja kuantitatif, keunggulan utama algoritma C4.5 terletak pada interpretabilitas model, di mana aturan keputusan yang dihasilkan mudah dipahami dan dapat dijadikan dasar pertimbangan kebijakan oleh pemangku kepentingan. Implikasi dari temuan ini adalah bahwa pendekatan data mining, khususnya Decision Tree C4.5, dapat dimanfaatkan sebagai sistem pendukung keputusan dalam program bantuan sosial untuk meningkatkan transparansi, akuntabilitas, dan efisiensi pengelolaan anggaran publik. Kontribusi penelitian ini terhadap body of knowledge terletak pada penguatan bukti empiris mengenai relevansi algoritma C4.5 dalam konteks klasifikasi sosial-ekonomi serta integrasi antara evaluasi kinerja model dan interpretasi kebijakan. Penelitian selanjutnya disarankan untuk mengeksplorasi

perbandingan dengan algoritma lain atau pendekatan ensemble, serta menguji model pada dataset yang lebih besar dan beragam wilayah guna meningkatkan generalisasi dan ketahanan model dalam implementasi nyata.

## REFERENCES

- [1] D. Purbaningrum, Hendri Hermawan Adinugraha, and M. Z. Piliang, "Transformasi Kebijakan Publik Menyiasati Ketidaktepatan Sasaran Bantuan Sosial Di Desa Pringsurat Untuk Kesejahteraan Masyarakat Yang Lebih Baik," *J. Pembang. dan Kebijak. Publik*, vol. 15, no. 2, pp. 262–284, 2024, doi: 10.36624/jpkp.v15i2.116.
- [2] Tasya Aulia Putri, Akhmad Andri Firdaus, Marwah Syifa Ardila, and Ahmad Faiz Ibrahim, "Indikator dan Permasalahan Makro Ekonomi," *JEMBA J. Econ. Bus. Account.*, vol. 1, no. 3, pp. 149–174, 2023, doi: 10.54783/jemba.v1i3.63.
- [3] R. Indawati, "Efektivitas Penyaluran Program Bantuan Sosial Tunai (Bst) Di Masa Pandemi Covid-19 Di Kelurahan Kalabirang Kecamatan Pattalassang Kabupaten Takalar)," *J. Adm. Negara*, vol. 28, no. 1, pp. 24–41, 2022, doi: 10.33509/jan.v28i1.1673.
- [4] P. Rahayu and Hanik Fitriani, "Efektivitas Program Bantuan Pangan Non Tunai Dalam Pengentasan Kemiskinan Di Kabupaten Ponorogo," *Niqosiya J. Econ. Bus. Res.*, vol. 3, no. 1, pp. 146–155, 2023, doi: 10.21154/niqosiya.v3i1.1758.
- [5] A. Rahmawati1, Muhammad Win Afgani2, "Evaluasi Program Peningkatan Kompensasi Guru Pendidikan Agama Islam untuk Meningkatkan Mutu Guru PAI di Kota Palembang," *Indones. Res. J. Educ. Web*, vol. 5, no. 3, pp. 649 – 657, 2025.
- [6] S. Y. Mozin, S. Abdullah, and N. Sawali, "Pemanfaatan Teknologi Cerdas Untuk Pelayanan Publik: Study Tentang e-Government Dan Smart City Berbasis ICT Big Data Dan AI," *JPS J. Publ. Stud.*, vol. 2, no. 2, pp. 117–130, 2025, [Online]. Available: <https://jurnal.fisip.untad.ac.id/index.php/jps/article/view/1948>
- [7] D. M. Musa *et al.*, "Penerapan Data Mining Untuk Klasifikasi Data Penjualan Pakan Ternak Terlaris Dengan Algoritma C4.5," *J. Teknol. Inform. dan Komput.*, vol. 10, no. 1, pp. 172–186, 2024, doi: 10.37012/jtik.v10i1.1985.
- [8] Ardiansyah Risiko Anwari and S. Sukirman, "Recommendation System To Select a Major of Vocational School Using Decision Tree," *J. Tek. Inform.*, vol. 5, no. 2, pp. 589–598, 2024, doi: 10.52436/1.jutif.2024.5.2.1327.
- [9] A. Rahman, "Praktek Permodelan Kebijakan," *AMU Press*, vol. 1, no. 1, pp. 1–77, 2025, [Online]. Available: <https://ejournal.amertamedia.co.id/index.php/press/article/view/392>
- [10] Abd. Muiz, R. Anisah, U. Khoiruddin, and E. Indrioko, "Kebijakan Pendidikan Dalam Mengatasi Masalah Kualitas, Kuantitas Efektivitas dan Efisiensi," *J. IHSAN J. Pendidik. Islam*, vol. 2, no. 3, pp. 46–64, 2024, doi: 10.61104/ihsan.v2i3.272.
- [11] M. Hasanuddin, B. E. Susanto, S. Ginting, and F. Rizaldi, "Analisis Minat Siswa Kelas 1 SMK Pada Ekstrakurikuler Sepak Bola Dengan Metode Technology Acceptance Model," *J. Komput. Teknol. Inf. Sist. Inf.*, vol. 4, no. 1, pp. 52–58, 2025, doi: 10.62712/juktisi.v4i1.344.
- [12] Jamiah Nurhakiki and Yahfizham Yahfizham, "Studi Kepustakaan: Pengenalan 4 Algoritma Pada Pembelajaran Deep Learning Beserta Implikasinya," *Pendek. J. Pendidik. Berkarakter*, vol. 2, no. 1, pp. 270–281, 2024, doi: 10.51903/pendekar.v2i1.598.
- [13] P. Murtopo, I. Dwi Yulianto, S. Suparno, and S. Saparuddin, "Penerapan Pemodelan Konsep Dinamis dalam Keputusan Bisnis: Optimalisasi Keputusan dengan Linear Optimization, Decision Tree dan Scenario Test," *J. Account. Financ. Manag.*, vol. 6, no. 2, pp. 475–484, 2025, doi: 10.38035/jafm.v6i2.1815.
- [14] C. Saskia Rafika, R. Maliq Reynanda, and A. Puspita Sari, "Penerapan Decision Tree CART untuk Klasifikasi Risiko Gagal Studi Mahasiswa," *J. Mhs. Tek. Inform.*, vol. 4, no. 2, pp. 37–46, 2025, doi: 10.35473/jamastika.v4i2.4145.
- [15] O. P. Moerdyanto and I. K. D. Nuryana, "Prediksi Kelulusan Tepat Waktu Menggunakan Pendekatan Pohon Keputusan Algoritma Decision Tree," *J. Informatics Comput. Sci.*, vol. 05, no. 1, pp. 90–96, 2023, [Online]. Available: <https://ejournal.unesa.ac.id/index.php/jinacs/article/view/55329>
- [16] P. Ramadani, R. Fadillah, Q. Adawiyah, and B. R. Al Ghazali, "Perbandingan Algoritma Naïve Bayes, C4.5, dan K-Nearest Neighbor untuk Klasifikasi Kelayakan Program Keluarga Harapan," *J. Media Inform.*, vol. 6, no. 2, pp. 775–782, 2024.