VOL. 02 NO. 1 JANUARY 2025
P-ISSN: XXXX-XXXX | E-ISSN: XXXX-XXXX
DOI: XX.XXXXX/jcsai.vXXX.XXXX

JOURNAL OF COMPUTER SCIENCE AND ARTIFICIAL INTELLIGENCE (JCSAI)

IMPLEMENTASI ALGORITMA C4.5 UNTUK MENINGKATKAN AKURASI KLASIFIKASI PENERIMA BANTUAN SOSIAL DI INDRAMAYU

Nuzlatul Hidayati[1]; Nana Suarna[2]; Irfan Ali[3]; Dodi Solihudin[4]

Program Studi Teknik Informatika[124]
Program Studi Rekayasa Perangkat Lunak[3]
STMIK IKMI CIREBON
nuzlatulhidayati02@gmail.com[1], nana.ikmi@gmail.com[2] ,
irfanali0.0@gmail.com[3],
dodisolihudin@gmail.com[4]

(*) Corresponding Author : nuzlatulhidayati02@gmail.com
Received: date; Accepted: date; Published: date

Abstract— *This study aims to improve the classification accuracy of social assistance eligibility in Indramayu using the C4.5 algorithm. Inefficient manual selection often results in social assistance programs being misallocated. The C4.5 algorithm is applied to develop a decision tree-based classification model. The dataset consists of 392 records with 14 attributes, including income, number of dependents, age, and occupation. Data processing is conducted using RapidMiner, with a 70:30 split for training and testing. The research follows the Knowledge Discovery in Databases (KDD) stages, including data selection, cleaning, and model evaluation using a confusion matrix. The results show that the classification model achieves an accuracy of 93.22%, with a precision of 92.34% for the "eligible" category and 100% for the "ineligible" category. The most significant attributes are the number of dependents, the number of working family members, and age. Key classification rules indicate that having more than 2.5 dependents is a decisive factor for eligibility. Attribute weight visualization highlights that the number of dependents (0.429) has the greatest influence, followed by working family members (0.302) and age (0.269). This study demonstrates the effectiveness of the C4.5 algorithm in improving the accuracy of social assistance eligibility classification. The model is expected to support more efficient and targeted assistance distribution. Additionally, this research offers opportunities for further development, such as implementation in other regions with larger datasets to enhance model reliability.*

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

Intisari— Penelitian ini bertujuan meningkatkan akurasi klasifikasi kelayakan penerima bantuan sosial di Indramayu menggunakan algoritma C4.5. Seleksi manual yang tidak efisien sering menyebabkan program bantuan sosial tidak tepat sasaran. Algoritma C4.5 diterapkan untuk membangun model klasifikasi berbasis pohon keputusan. Dataset terdiri dari 392 data penerima dengan 14 atribut, seperti penghasilan, jumlah tanggungan, usia, dan pekerjaan. Data diolah menggunakan RapidMiner dengan pembagian 70% untuk pelatihan dan 30% untuk pengujian. Proses penelitian mengikuti tahapan Knowledge Discovery in Databases (KDD), meliputi seleksi, pembersihan data, dan evaluasi model menggunakan confusion matrix. Hasilnya, model klasifikasi mencapai akurasi 93,22%, dengan precision kategori "layak" sebesar 92,34% dan "tidak layak" 100%. Atribut paling signifikan adalah jumlah tanggungan, anggota keluarga bekerja, dan usia. Aturan utama klasifikasi menunjukkan bahwa jumlah tanggungan lebih dari 2,5 menjadi faktor penentu kelayakan. Visualisasi bobot atribut menempatkan jumlah tanggungan (0,429) sebagai pengaruh terbesar, diikuti anggota keluarga bekerja (0,302) dan usia (0,269). Penelitian ini membuktikan efektivitas algoritma C4.5 dalam meningkatkan akurasi klasifikasi penerima bantuan sosial. Model ini diharapkan mendukung distribusi bantuan yang lebih tepat sasaran. Penelitian ini juga membuka peluang pengembangan, seperti penerapan di wilayah lain dengan dataset lebih besar untuk meningkatkan keandalan model.

Kata Kunci: *Algoritma C4.5, Klasifikasi Data, Bantuan Sosial.*

VOL. 02 NO. 1 JANUARY 2025
P-ISSN: XXXX-XXXX | E-ISSN: XXXX-XXXX
DOI: XX.XXXXX/jcsai.vXXX.XXXX

INTRODUCTION

Kemajuan pesat dalam informatika berdampak besar pada berbagai aspek kehidupan, termasuk teknologi, bisnis dan pendidikan. Inovasi ini tidak hanya meningkatkan efisiensi operasional, tetapi juga memungkinkan solusi yang lebih canggih untuk masalah sosial, seperti penggunaan data mining dan algoritma klasifikasi dalam analisis data. Di sektor pelayanan publik, kemampuan untuk mengolah data dengan cepat dan akurat sangat penting, terutama dalam menentukan kelayakan penerima bantuan sosial.

Meskipun kemajuan teknologi pesat, tantangan besar tetap ada dalam pengelolaan data sosial, terutama dalam menentukan kelayakan penerima bantuan. Data yang besar dan kompleks menyulitkan analisis manual, terutama di daerah dengan kemiskinan tinggi seperti Kabupaten Indramayu. Penentuan kelayakan yang tidak tepat dapat menyebabkan bantuan tidak tepat sasaran, sehingga dibutuhkan teknologi, seperti algoritma, untuk memperbaiki akurasi dan efisiensi dalam pengambilan keputusan.

Beberapa penelitian sebelumnya telah menggunakan algoritma C4.5 untuk menentukan kelayakan penerima bantuan sosial. Penelitian oleh [1] memanfatkan Algoritma C4.5 untuk menentukan kelayakan penerima bantuan Covid-19. Sehingga, Algoritma C4.5 digunakan untuk membuat aturan pohon keputusan yang tepat guna memastikan penargetan bantuan yang akurat. Atribut yang paling berpengaruh dalam menentukan kelayakan adalah jumlah penghasilan, diikuti oleh jumlah tanggungan. Hal ini terlihat dari diagram pohon keputusan yang menunjukkan bahwa jumlah penghasilan menjadi simpul akar *(root node)* dan jumlah tanggungan menjadi faktor kedua dalam menentukan kelayakan penerima bantuan Covid-19. Penelitian [2] menggunakan algoritma C4.5 dalam klasifikasi penerima bantuan pangan non-tunai (BPNT) di Lampung Utara, yang menghasilkan akurasi prediksi hingga 98,5%. Hasil ini menunjukkan bahwa algoritma C4.5 dapat digunakan secara efektif untuk klasifikasi penerima bantuan dengan atribut tertentu, seperti penghasilan dan status rumah. Namun, keakuratan tinggi ini menunjukkan adanya potensi *overfitting* terhadap dataset tertentu, sehingga masih diperlukan validasi pada dataset yang lebih variatif untuk memastikan generalisasi yang lebih baik. Penelitian oleh [3]menerapkan algoritma CART untuk menentukan kelayakan penerima Program Keluarga Harapan (PKH). Algoritma CART, yang memiliki kemampuan klasifikasi multikategori, digunakan untuk menghasilkan tiga klasifikasi utama yaitu "layak","dipertimbangkan" dan"tidak layak". Meskipun memiliki fleksibilitas dalam klasifikasi, pendekatan ini masih dipengaruhi oleh bias subjektif dalam proses musyawarah awal, sehingga mengurangi objektivitas dalam pengambilan keputusan.

Penelitian ini bertujuan untuk menggunakan algoritma C4.5 dalam menentukan kelayakan penerima bantuan sosial di Indramayu, dengan harapan dapat meningkatkan akurasi dan efektivitas penyaluran bantuan yang lebih tepat sasaran. Diharapkan hasil penelitian ini akan memberikan manfaat besar bagi masyarakat dan membantu perangkat desa membuat keputusan yang lebih adil dan berbasis data valid dalam mendistribusikan bantuan sosial.

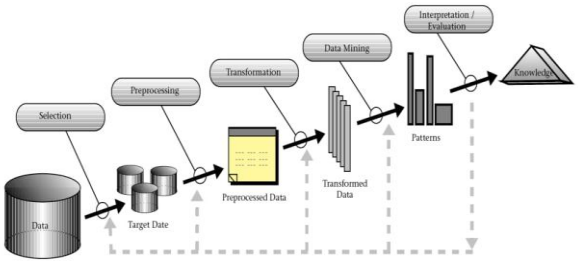
Penelitian ini menggunakan algoritma Decision Tree C4.5 untuk membangun model klasifikasi berdasarkan atribut-atribut penting seperti nama, jenis kelamin, usia, pekerjaan, jumlah tanggungan, penghasilan, dan status kelayakan. Dataset yang digunakan berasal dari aparat desa di Indramayu dan dianalisis menggunakan software RapidMiner. Diharapkan model yang dihasilkan dapat memberikan hasil yang akurat dan dapat diandalkan untuk menentukan kelayakan penerima bantuan sosial.

Hasil penelitian ini diharapkan dapat memperkuat penerapan teknologi C4.5 dalam menentukan kelayakan penerima bantuan sosial secara lebih adil dan tepat sasaran. Selain itu, penelitian ini dapat menjadi referensi bagi desa-desa lain dalam mengimplementasikan sistem serupa, serta meningkatkan peran informatika dalam kebijakan sosial yang lebih efektif dan adil. Dengan demikian, penyaluran bantuan sosial dapat menjadi lebih tepat sasaran dan berdasarkan data yang valid.

MATERIALS AND METHODS

Metode penelitian yang digunakan dalam penelitian ini adalah *Knowledge Discovery in Database* (KDD), yaitu sebuah metode yang digunakan untuk mengekstraksi data guna memperoleh informasi atau pola tertentu dari data yang telah dipilih sebelumnya. Proses dalam metode KDD mencakup beberapa tahapan utama, yaitu pemilihan data *(Data Selection)*, praproses data *(Preprocessing)*, transformasi data *(Transformation),* penggalian data *(Data Mining),* dan evaluasi hasil *(Evaluation),* yang semuanya

saling berkaitan untuk menghasilkan temuan yang akurat dan relevan[4].



Gambar. 1 Proses KDD

Pada gambar 1 terlihat sebuah metode KDD atau tahapan tahapan yang dilakukan dari peneliian ini mulai dari Data Selection, Preprocessing, Data Transformation, Data Mining, Evaluasi.

Data Selection

Langkah pertama ini merupakan langkah yang sangat penting dalam memastikan bahwa data yang akan digunakan dalam proses Knowledge Discovery in Databases (KDD) adalah data yang tepat, relevan, dan berkualitas tinggi [5]. Data yang digunakan yaitu data penduduk Desa di Kabupaten Indramayu pada tahun 2024 berjumlah 392 dan ada 14 atribut seperti nama, jenis kelamin, tempat lahir, usia, status, hubungan keluarga, pendidikan, pekerjaan, jumlah tanggungan, penghasilan, jenis pekerjaan pasangan, jumlah anggota keluarga yang bekerja, dan status kelayakan(layak/tidak layak) untuk dijadikan kriteria prediksi kelayakan penerima bantuan sosial pada penelitian ini.

Preprocessing

Pada tahap ini, diperlukan proses dasar untuk membersihkan data, seperti menghilangkan noise. Sebelum proses data mining dimulai, Proses cleaning harus dilakukan pada data yang menjadi fokus utama dalam Knowledge Discovery in Database (KDD). Tahapan ini mencakup berbagai langkah, seperti menghapus data yang terduplikasi, memeriksa adanya inkonsistensi pada data, serta memperbaiki kesalahan, termasuk kesalahan pengetikan atau typo [6].

Data Transformation

Tahap transformasi data melibatkan penyesuaian kategori data agar sesuai dengan jenis data yang dibutuhkan. Proses ini mencakup pengubahan format, skala, atau struktur data sehingga dapat memenuhi kebutuhan analisis atau model yang akan diterapkan [7]. Hal ini bertujuan untuk mempermudah proses analisis di tahap data mining. Misalnya, atribut numerik dapat dinormalisasi untuk memastikan nilainya berada dalam rentang tertentu, atau atribut kategorikal dapat diubah menjadi bentuk numerik agar lebih mudah diolah oleh algoritma. Namun, dalam penelitian ini, tahap transformasi tidak dilakukan karena data sudah tersedia dalam format yang sesuai untuk digunakan langsung dalam proses klasifikasi menggunakan algoritma C4.5.

Data Mining

Pada tahap ini, Ada beberapa proses data mining pada penelitian ini yaitu: Melakukan penerapan *Split Data* yang berarti pembagian dataset menjadi data *training* dan data *testing*, dengan pembagian rasio 7:3. Selanjutnya, Melakukan penerapan algoritma *Decision Tree* yaitu metode klasifikasi yang digunakan untuk meningkatkan akurasi model dalam menentukan kelayakan penerima bantuan sosial [8]. Selanjutnya tambahkan operator *Apply model* yang berfungsi untuk menerapkan data guna menghasilkan sebuah klasifikasi atau prediksi pada dataset [9].

Evaluation

Pada tahapan akhir ini, pada Tools RapidMiner, Operator Performance dapat digunakan untuk menghasilkan output yang memungkinkan pola informasi dari hasil data mining dapat dipahami dengan jelas [6]. Pada langkah ini, hasil dari teknik data mining dievaluasi melalui pola representasi dan model prediksi untuk menilai apakah hipotesis yang ditetapkan telah tercapai. Tahapan ini juga berfungsi sebagai umpan balik untuk menyempurnakan proses data mining agar lebih optimal, atau menerima hasil yang mungkin tidak sesuai ekspektasi tetapi tetap bermanfaat.

RESULTS AND DISCUSSION

Data Selection

Pada tahap awal dalam proses data selection di aplikasi RapidMiner, digunakan fitur Read Excel, yang merupakan salah satu operator penting dalam mengimpor data dari file Excel. Fitur ini dirancang untuk membaca dan memproses file dengan format ".xlsx" atau ".xls" secara langsung [10]. Dengan menggunakan operator ini, seluruh data yang terdapat dalam file Excel akan diambil secara menyeluruh untuk dipersiapkan dan diproses lebih lanjut pada tahap-tahap analisis berikutnya. Langkah ini sangat penting karena memastikan bahwa semua data relevan telah berhasil dimuat ke dalam sistem secara lengkap dan akurat [11].

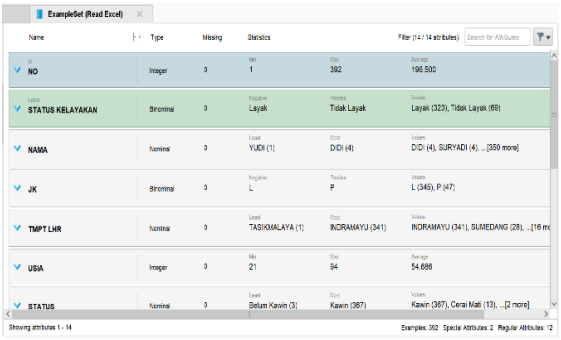
VOL. 02 NO. 1 JANUARY 2025
P-ISSN: XXXX-XXXX | E-ISSN: XXXX-XXXX
DOI: XX.XXXXX/jcsai.vXXX.XXXX

Tabel. 1 Data

No	Nama	JK	Tempat Lahir	..	Status Kelayakan
1	Rasta	L	Indramayu	...	Layak
2	Tata Satriaman	L	Indramayu	...	Layak
3	Wawan Rastiwan	L	Indramayu	...	Layak
4	Uya	L	Indramayu	...	Layak
5	Tarsoma	L	Indramayu	...	Tidak Layak
...	...	...	...	...	...
...	...	...	...	...	...
392	Isna Sumarna	L	Indramayu	...	Tidak Layak

Preprocessing

Pada tahap ini dilakukan *pre-processing* data, yaitu proses persiapan sebelum data diolah dalam prosess data mining. *Pre-processing* bertujuan untuk mengurangi ataupun menghilangkan data yang memiliki nilai yang hilang (missing value). Selanjutnya, pada tahap pembersihan (cleaning), dilakukan penghapusan dan penyesuaian terhadap data yang tidak lengkap atau duplikat. Proses ini bertujuan untuk mengurangi jumlah data yang tidak relevan.

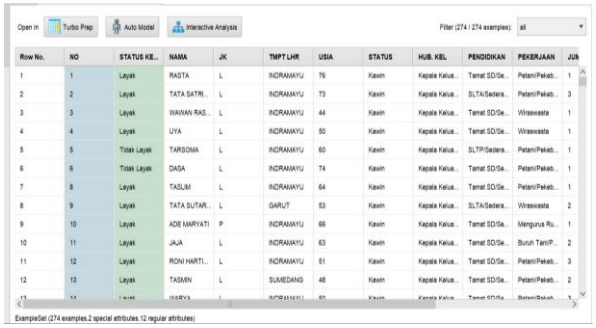


Gambar. 2 Preprocessing

Gambar di atas menunjukkan hasil dari proses pengecekan dan pembersihan data untuk memastikan tidak ada kekosongan atau nilai yang hilang pada dataset. Hasilnya menunjukkan bahwa seluruh data telah terisi dengan lengkap, sehingga jumlah data penerima bantuan sosial tetap sebanyak 392 data tanpa adanya perubahan.

Data Transformation

Pada tahap *Transformasi* biasanya dilakukan penyelarasan data seperti normalisasi atau pengaturan format agar data dapat digunakan secara optimal dalam proses data mining. Biasanya, langkah ini diperlukan untuk memastikan setiap atribut data berada pada skala atau format yang seragam. Namun, dalam penelitian ini penulis tidak melakukan proses *transformasi* karena data sudah siap untuk dianalisis tanpa perlu perubahan tambahan.

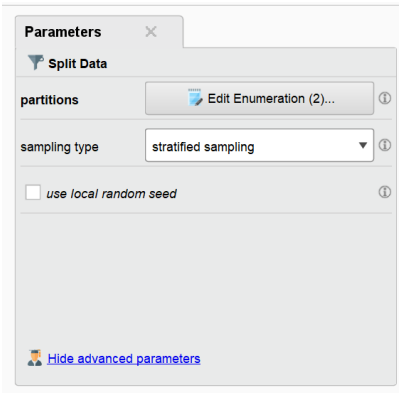


Gambar. 3 Transformation

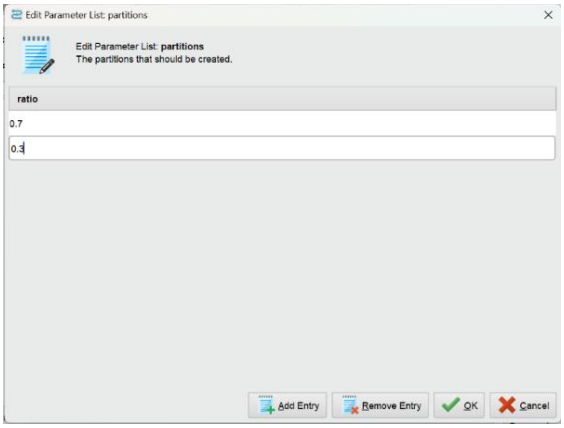
Data Mining

a. Split Data

Langkah awal pada tahap proses *Data Mining* yaitu penggunaan operator Split Data berfungsi untuk membagi dataset menjadi dua bagian sesuai ratio , yaitu data training untuk melatih model dan data testing untuk menguji kinerjanya dengan ratio pembagian yang ditentukan.Selanjutnya, pada parameters *Split* terdapat menu Edit Enumeration. Pada menu *Partitions*, data dibagi menjadi dua bagian dengan ratio 7:3, dimana 70% digunakan sebagai data training dan 30% sebagai *data testing*. Pada menu *Sampling Type*, pilih opsi *Stratified Sampling* digunakan untuk memastikan setiap kategori atau atribut dalam dataset tetap seimbang disetiap bagian, sehingga hasil analisis lebih representatif dan akurat.



Gambar. 4 Parameter Split Data



Gambar. 5 Edit Parameters Partitions

b. Decision Tree

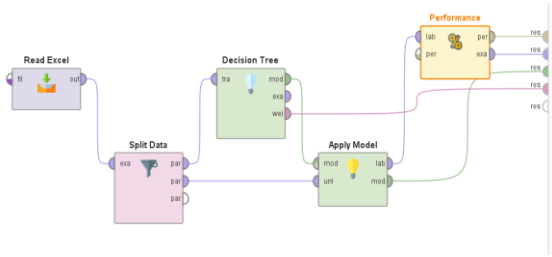
Metode yang digunakan dalam penelitian ini adalah Algoritma C4.5, yang bertujuan untuk melakukan klasifikasi kelayakan penerima bantuan sosial. Algoritma ini dipilih karena kemampuannya yang baik dalam menangani data yang kompleks dan menghasilkan model yang mudah dipahami. Setelah data diproses pada tahap sebelumnya, langkah selanjutnya adalah mengolah data tersebut menggunakan operator Decision Tree di dalam perangkat lunak. Operator ini akan membentuk pohon keputusan berdasarkan algoritma C4.5, yang memanfaatkan atribut-atribut yang ada pada dataset untuk menilai kelayakan seseorang dalam menerima bantuan sosial. Algoritma C4.5 cocok digunakan dalam penelitian ini karena mampu menangani berbagai tipe data serta menghasilkan aturan klasifikasi yang jelas, sehingga memudahkan dalam mengevaluasi hasil model yang terbentuk. Model yang dihasilkan dari algoritma ini diharapkan dapat membantu mempermudah proses pengambilan keputusan terkait kelayakan penerima bantuan sosial.

c. Apply Model

Langkah berikutnya adalah menambahkan operator *Apply Model* dalam proses. Operator ini berfungsi untuk menerapkan model yang sudah dilatih (*training*) sebelumnya pada dataset uji. Dengan menggunakan operator ini, model yang sudah dipelajari akan diaplikasikan pada data baru yang belum pernah dianalisis oleh model. Tujuan utamanya adalah untuk mengukur akurasi model terhadap data uji tersebut. Operator *Apply Model* memungkinkan model untuk menerapkan pola dan aturan yang telah dipelajari selama tahap pelatihan, sehingga dapat diketahui seberapa baik model mampu mengklasifikasikan data secara akurat. Hasil dari proses ini tidak hanya berupa prediksi label, tetapi juga digunakan untuk menghitung tingkat akurasi model, yang menjadi indikator seberapa baik model bekerja dalam mengklasifikasikan data uji secara keseluruhan.

d. Evaluasi

Pada tahapan Evaluasi, penelitian ini menggunakan operator *Performance* di *tools RapiMiner* untuk menilai kinerja algoritma yang akan diuji. Proses evaluasi dimulai dengan melatih model data training dengan menggunakan metode Decision Tree Algoritma C4.5. Setelah model terbentuk, data yang sama diuji kembali terhadap model tersebut untuk mengukur tingkat akurasi hasil klasifikasi. Setelah seluruh operator digunakan, model yang dihasilkan bisa dilihat pada Gambar 7 dibawah ini:

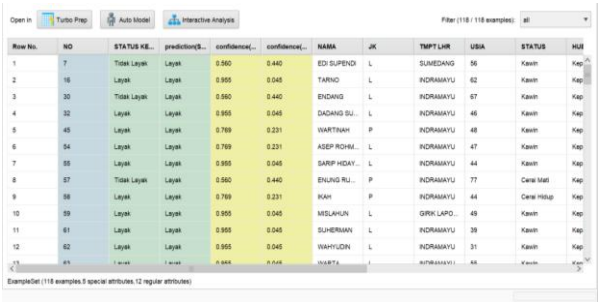


Gambar. 6 Hasil Pemodelan Decision Tree

Berdasarkan hasil penelitian data mining menggunakan algoritma C4.5, proses klasifikasi kelayakan penerima bantuan sosial telah diterapkan dengan baik. Dengan hasil perhitungan yang dilakukan menggunakan *RapidMiner*, dapat dianalisis tingkat akurasi model yang dihasilkan. Berikut ditampilkan pada gambar berikut:

1. Example Data

a. Data



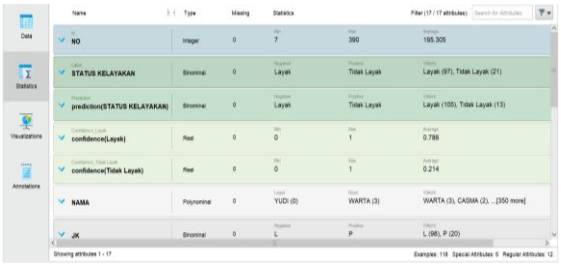
Gambar. 7 Data

Gambar 7 menunjukkan 118 data uji yang digunakan untuk menguji model klasifikasi kelayakan penerima bantuan sosial. Dataset ini mencakup 5 atribut khusus, seperti ID, status kelayakan, dan prediksi kelayakan, serta 12 atribut reguler, termasuk jumlah tanggungan, usia, dan penghasilan. Atribut khusus berfungsi sebagai elemen utama dalam proses klasifikasi, sedangkan

VOL. 02 NO. 1 JANUARY 2025
P-ISSN: XXXX-XXXX | E-ISSN: XXXX-XXXX
DOI: XX.XXXXX/jcsai.vXXX.XXXX

atribut reguler memberikan konteks tambahan untuk mendukung analisis.

b. Statistic

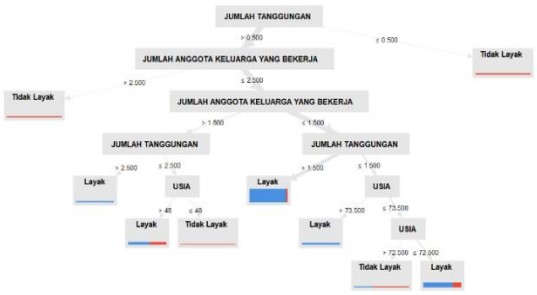


Gambar. 8 Statistic

Berdasarkan Gambar 8, data uji terdiri dari 17 atribut, yaitu 5 atribut khusus dan 12 atribut reguler. Atribut khusus, seperti ID, status kelayakan, prediksi, dan *confidence*, digunakan untuk mengevaluasi model klasifikasi yang dihasilkan algoritma C4.5. Atribut reguler mendukung analisis dalam proses pengambilan keputusan.

2. Decision Tree

a. Tree



Gambar. 9 Decision Tree

Gambar 9 menunjukkan pohon keputusan yang menghasilkan aturan klasifikasi. Contoh aturan:

1. Jika *Jumlah Tanggungan* ≤ 0,5, maka status *Tidak Layak*.
2. Jika *Jumlah Tanggungan* > 0,5 dan *Jumlah Anggota Keluarga yang Bekerja* > 2,5, maka *Tidak Layak*.
3. Jika *Jumlah Tanggungan* > 0,5 dan *Jumlah Anggota Keluarga yang Bekerja* ≤ 2,5, maka kelayakan bergantung pada usia dan jumlah tanggungan.

Proses ini menunjukkan bagaimana algoritma C4.5 menggunakan atribut utama secara bertahap untuk menghasilkan klasifikasi akurat.

b. Decision Tree

Tree

```
JUMLAH TANGGUNGAN > 0.500
|   JUMLAH ANGGOTA KELUARGA YANG BEKERJA > 2.500: Tidak Layak {Layak=0, Tidak Layak=5}
|   JUMLAH ANGGOTA KELUARGA YANG BEKERJA ≤ 2.500
|   |   JUMLAH ANGGOTA KELUARGA YANG BEKERJA > 1.500
|   |   |   JUMLAH TANGGUNGAN > 2.500: Layak {Layak=7, Tidak Layak=0}
|   |   |   JUMLAH TANGGUNGAN ≤ 2.500
|   |   |   |   USIA > 48: Layak {Layak=14, Tidak Layak=11}
|   |   |   |   USIA ≤ 48: Tidak Layak {Layak=0, Tidak Layak=2}
|   |   JUMLAH ANGGOTA KELUARGA YANG BEKERJA ≤ 1.500
|   |   |   JUMLAH TANGGUNGAN > 1.500: Layak {Layak=147, Tidak Layak=7}
|   |   |   JUMLAH TANGGUNGAN ≤ 1.500
|   |   |   |   USIA > 73.500: Layak {Layak=17, Tidak Layak=0}
|   |   |   |   USIA ≤ 73.500
|   |   |   |   |   USIA > 72.500: Tidak Layak {Layak=1, Tidak Layak=2}
|   |   |   |   |   USIA ≤ 72.500: Layak {Layak=40, Tidak Layak=12}
JUMLAH TANGGUNGAN ≤ 0.500: Tidak Layak {Layak=0, Tidak Layak=9}
```

Gambar. 10 Decision Tree

Berdasarkan Gambar 10, hasil pohon keputusan menunjukkan bahwa model klasifikasi berhasil mengidentifikasi aturan utama yang menentukan kelayakan penerima bantuan sosial. Setiap aturan mencerminkan pengaruh variabel-variabel utama, seperti jumlah tanggungan, usia, dan jumlah anggota keluarga yang bekerja, dalam proses pengambilan keputusan. Dengan memanfaatkan aturan ini, model dapat memberikan penjelasan yang transparan mengenai dasar klasifikasi untuk setiap data penerima bantuan sosial.

3. Attribute Weights

a. Data

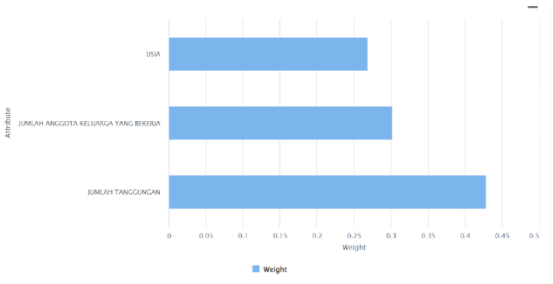
attribute	weight
USIA	0.269
JUMLAH TANGGUNGAN	0.429
JUMLAH ANGGOTA KELUARGA YANG BEKERJA	0.302

Gambar. 11 Attribute Weight

Hasil analisis seperti yang ditunjukkan pada Gambar 11 mengungkapkan bahwa atribut Jumlah Tanggungan memiliki pengaruh paling besar dalam menentukan kelayakan penerima bantuan sosial, dengan bobot sebesar 0.429. Atribut berikutnya adalah Jumlah Anggota Keluarga yang Bekerja dengan bobot 0.302, diikuti oleh Usia dengan bobot 0.269. Temuan ini menunjukkan bahwa variabel jumlah tanggungan memiliki peran paling signifikan dalam klasifikasi kelayakan, sementara atribut

lainnya memberikan kontribusi pendukung yang penting.

b. Weight Visualizations

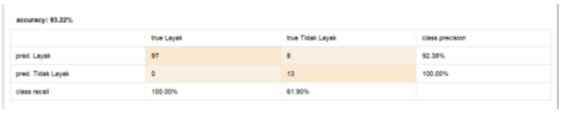


Gambar. 12 Weight Visualizations

Gambar 12 menunjukkan visualisasi bobot dari setiap atribut yang digunakan dalam klasifikasi. Berdasarkan hasil ini, dapat dilihat bahwa jumlah tanggungan adalah faktor yang paling berpengaruh dalam menentukan kelayakan penerima bantuan sosial, dengan bobot tertinggi. Sementara itu, jumlah anggota keluarga yang bekerja dan usia juga memainkan peran penting, meskipun dengan bobot yang lebih rendah. Visualisasi ini memudahkan untuk memahami bagaimana setiap atribut berkontribusi dalam proses pengambilan keputusan, sehingga kita bisa melihat dengan jelas faktor mana yang paling menentukan dalam klasifikasi.

4. Performance Vector

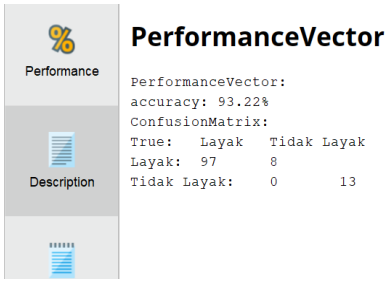
a. Performance



Gambar. 13 Hasil Akurasi Decision Tree

Berdasarkan gambar 13, model klasifikasi yang dibangun menggunakan algoritma C4.5 menghasilkan tingkat akurasi yang sangat tinggi yaitu sebesar 93.22%, *Class precission* menunjukkan hasil prediksi untuk kategori Layak sebesar 92.34%, sedangkan untuk kategori Tidak layak 100.00%. Hal ini menunjukan bahwa algoritma *Decision Tree* sangat efektif diterapkan pada dataset ini karena mampu menghasilkan keputusan klasifikasi yang sangat akurat.

b. Description



Gambar. 14 Description Performance

Berdasarkan gambar 14 menunjukkan detail hasil evaluasi model, dimana model klasifikasi menghasilkan 97 data *true* Layak, 13 data *true* Tidak Layak dan 8 data *false* Layak. Hasil ini menunjukkan bahwa model C4.5 dapat mengklasifikasikan penerima bantuan sosial dengan cukup akurat, sehingga hal ini mendukung pengambilan keputusan yang lebih akurat dan efisien.

CONCLUSION

Penelitian ini menyimpulkan bahwa jumlah tanggungan merupakan kriteria paling berpengaruh dalam menentukan kelayakan penerima bantuan sosial, diikuti oleh jumlah anggota keluarga yang bekerja dan usia penerima. Proses klasifikasi menggunakan algoritma C4.5 berhasil menghasilkan pohon keputusan yang mempermudah identifikasi penerima bantuan yang layak dengan tingkat akurasi yang tinggi. Dengan tingkat akurasi mencapai 93,22%, algoritma C4.5 terbukti sangat efektif diterapkan pada dataset bantuan sosial di Indramayu.

Untuk meningkatkan keakuratan, disarankan menambahkan atribut relevan lainnya, seperti kondisi rumah atau status kepemilikan tanah. Penelitian ini juga dapat dikembangkan lebih lanjut dengan membandingkan algoritma C4.5 dengan metode klasifikasi lainnya, seperti Random Forest atau Support Vector Machine. Pemerintah daerah diharapkan dapat mengadopsi teknologi berbasis data mining dalam proses pengambilan keputusan untuk memastikan distribusi bantuan sosial lebih tepat sasaran. Penelitian lanjutan juga disarankan menggunakan dataset yang lebih besar dan bervariasi untuk menguji generalisasi dan keandalan model.

VOL. 02 NO. 1 JANUARY 2025
P-ISSN: XXXX-XXXX | E-ISSN: XXXX-XXXX
DOI: XX.XXXXX/jcsai.vXXX.XXXX

REFERENCE

[1] A. Riadi, S. R. Andani, and R. R. Dewi, "Data Mining Algoritma C4.5 menentukan Kelayakan Penerima Bantuan Covid 19," *JUKI J. Komput. dan Inform.*, vol. 3, no. 2, pp. 44–51, 2021, doi: 10.53842/juki.v3i2.60.

[2] R. A. Islahudin, S. Rahmatullah, A. Afandi, and S. Safitri, "Algoritma C4.5 Untuk Memprediksi Kelayakan Penerima Bantuan Pangan Non Tunai," *J. Inform.*, vol. 22, no. 2, pp. 147–159, 2022, doi: 10.30873/ji.v22i2.3367.

[3] A. Rasyid, S. Gilbijatno, A. W. Pramudya, D. Prasetyo, and T. Informatika, "Implementasi Algoritma Decision Tree CART untuk Deteksi Dini Penyakit FLUTD pada Kucing," vol. 4, pp. 440–445, 2025.

[4] A. Triyono, A. Faqih, and G. Dwilestari, "Implementation of the Naive Bayes Method in Sentiment Analysis of Spotify Application Reviews," vol. 4, no. 2, 2025.

[5] A. Fakih, M. A. Hamzami, M. R. Hadianto, and N. I. Shafira, "Perbandingan Akurasi Algoritma C4 . 5 dan K-NN Untuk Prediksi," vol. 3, no. c, pp. 18–25, 2025.

[6] F. Firmansyah, N. Suarna, and W. Prihartono, "KLASIFIKASI HASIL PENJUALAN PAKAIAN MENGGUNAKAN ALGORITMA," *(Jurnal Mhs. Tek. Inform.*, vol. 9, no. 1, pp. 1291–1299, 2025.

[7] M. Ta and H. Ibrahimy, "PENERAPAN DECISION TREE C4 . 5 DALAM MEMPREDIKSI PREDIKAT TERBAIK DI," vol. 2, no. 1, pp. 61–68, 2025.

[8] L. Pasiolo *et al.*, "PENYAKIT STROKE DENGAN ALGORITMA SUPPORT," vol. 7, no. 1, pp. 61–73.

[9] U. Maltuf and Z. Fatah, "PENERAPAN ALGORITMA DICISION TREE UNTUK KLASIFIKASI KONSUMSI ENERGI LISRIK RUMAH TANGGA DENGAN PENGGUNAAN RAPID MINER," *J. Ilm. multidisiplin ilmu*, vol. 1, no. 2, pp. 99–105, 2024.

[10] R. Masdalipa, "Klasifikasi Data Penggunaan Obat Dengan Algoritma Naïve Bayes Di Rumah Sakit Umum Daerah (Rsud) Kota Pagar Alam JURNAL MEDIA INFORMATIKA [JUMIN]," vol. 7, no. 1, pp. 398–413, 2025.

[11] M. W. Martadiansyah, A. Ghufron, R. A. Hidayah, and D. Salzabila, "Perbandingan Algoritma C4 . 5 dengan Naive Bayes Untuk Klasifikasi Curah Hujan," vol. 3, no. c, pp. 8–17, 2025.