

IMPLEMENTASI ALGORITMA *DECISION TREE* UNTUK KLASIFIKASI KELAYAKAN DISTRIBUSI DANA BANTUAN SOSIAL

Hartana¹, Raden Deasy Mandasari²

¹Universitas Bina Sarana Informatika
*e-mail: hartana.han@bsi.ac.id

²Universitas Bina Sarana Informatika
e-mail: deasy.rde@bsi.ac.id

Abstrak

Penelitian ini mengkaji penggunaan algoritma pohon keputusan, khususnya C4.5, dalam klasifikasi kelayakan distribusi dana bantuan sosial di Kelurahan Bangka Belitung Laut, Kecamatan Pontianak Tenggara. Ketidaktepatan distribusi yang terjadi disebabkan oleh penyeleksian penerima berdasarkan rekomendasi perangkat desa yang hanya mempertimbangkan jenis pekerjaan. Untuk mengatasi masalah ini, penelitian ini mengembangkan sistem berbasis algoritma C4.5 yang mengolah data sensus ekonomi penduduk. Penelitian ini menggunakan metode kuantitatif dengan pendekatan deskriptif analitis, mengumpulkan data melalui studi pustaka dan sensus ekonomi, dan mengolah data menggunakan aplikasi Rapid Miner. Model klasifikasi yang dihasilkan dievaluasi menggunakan metode *10-fold cross-validation* untuk memastikan akurasi yang tinggi. Hasil penelitian menunjukkan bahwa algoritma C4.5 mencapai akurasi, *precision*, dan *recall* sebesar 100%. Model pohon keputusan yang dihasilkan menunjukkan bahwa atribut utama yang menentukan kelayakan adalah jenis pekerjaan dan pendapatan. Beberapa aturan yang dihasilkan dari model ini, seperti penduduk yang belum bekerja dengan pendapatan di bawah Rp29.500 layak menerima bantuan, memberikan panduan yang jelas bagi pengambil kebijakan. Implementasi algoritma ini diharapkan dapat meningkatkan keadilan dan efektivitas distribusi dana bantuan sosial di Kelurahan Bangka Belitung Laut, mengurangi ketidakpuasan masyarakat, dan mencegah potensi konflik sosial. Penelitian ini merekomendasikan adopsi model di wilayah lain dengan penyesuaian terhadap data lokal untuk meningkatkan keadilan distribusi bantuan secara lebih luas.

Kata Kunci: Algoritma C4.5, *Cross-Validation*, Distribusi Bantuan Sosial, Klasifikasi Kelayakan, Pohon Keputusan

Abstract

This study examined the use of decision tree algorithms, specifically C4.5, in classifying the eligibility for social assistance fund distribution in Kelurahan Bangka Belitung Laut, Kecamatan Pontianak Tenggara. Inaccuracies in distribution were caused by recipient selection based on village officials' recommendations, which only considered the type of occupation. To address this issue, this study developed a system based on the C4.5 algorithm to process economic census data. This research used a quantitative method with a descriptive analytical approach, collecting data through literature studies and economic census, and processing data using the Rapid Miner application. The classification model generated was evaluated using 10-fold cross-validation to ensure high accuracy. The results showed that the C4.5 algorithm achieved 100% accuracy, precision, and recall. The decision tree model indicated that the main attributes determining eligibility were occupation and income. Some rules derived from this model, such as those who are unemployed with an income below Rp29,500 being eligible for assistance, provided clear guidelines for policymakers. The implementation of this algorithm is expected to improve fairness and effectiveness in the distribution of social assistance funds in Kelurahan Bangka Belitung Laut, reduce public dissatisfaction, and prevent potential social conflicts. This study recommended

adopting the model in other areas with adjustments to local data to enhance broader fairness in aid distribution.

Keywords: *C4.5 Algorithm, Cross-Validation, Decision Tree, Eligibility Classification, Social Assistance Distribution*

1. Pendahuluan

Dalam pemberian bantuan sosial di Kelurahan Bangka Belitung Laut, Kecamatan Pontianak Tenggara, terdapat ketidaktepatan dalam distribusi dana bantuan. Hal ini disebabkan oleh penyeleksian penerima bantuan yang dilakukan berdasarkan rekomendasi perangkat desa, terutama RT, yang hanya mempertimbangkan jenis pekerjaan penduduk. Akibatnya, penduduk yang tergolong tidak mampu sering kali tidak menerima bantuan yang seharusnya mereka terima. Ketidaktepatan ini bukan hanya menyebabkan ketidakpuasan di kalangan masyarakat, tetapi juga berpotensi menimbulkan masalah sosial yang lebih luas jika tidak segera diatasi.

Rasional penelitian ini adalah pentingnya sistem yang lebih akurat dan objektif dalam menentukan kelayakan penerima bantuan sosial. Penelitian sebelumnya telah menunjukkan bahwa metode tradisional yang berbasis rekomendasi personal sering kali kurang akurat dan tidak adil. Menurut penelitian oleh (Abdulhafedh, 2022; Sugiono, Taufik, & Faizal Amir, 2020), algoritma klasifikasi seperti *Decision Tree* dapat memberikan akurasi yang lebih tinggi dalam berbagai aplikasi prediktif.

Urgensi dari penelitian ini terletak pada kebutuhan mendesak untuk memperbaiki sistem distribusi bantuan sosial di Kelurahan Bangka Belitung. Penelitian oleh (Ladianto, 2022) menunjukkan bahwa penerapan algoritma *data mining*, termasuk *Decision Tree*, dapat meningkatkan efektivitas dan efisiensi dalam penentuan penerima bantuan (Aulia, Suarna, & Prihartono, 2024; Nur, Rohim, Purnamasari, & Ali, 2024). Oleh karena itu, implementasi algoritma C4.5, yang merupakan varian dari *Decision Tree*, dipilih sebagai solusi potensial untuk mengatasi masalah ini.

Permasalahan utama dalam penelitian ini adalah bagaimana meningkatkan akurasi penentuan penerima bantuan sosial di Kelurahan Bangka Belitung. Alternatif solusi yang dipertimbangkan termasuk berbagai teknik

klasifikasi *data mining* seperti *Logistic Regression*, *Naive Bayes* (Ernamia & Herliana, 2022), dan *Support Vector Machines*. Namun, berdasarkan studi oleh (Ba'abbad, Althubiti, Alharbi, Alfarsi, & Rasheed, 2021), algoritma *Decision Tree*, khususnya C4.5, menunjukkan kinerja yang unggul dalam hal interpretabilitas dan akurasi pada *dataset* yang serupa.

Solusi yang dipilih dalam penelitian ini adalah implementasi algoritma C4.5 untuk klasifikasi kelayakan penerima bantuan sosial. Algoritma ini tidak hanya mampu menghasilkan model prediksi yang akurat tetapi juga mudah diinterpretasikan, sehingga dapat digunakan oleh pemerintah daerah dan *stakeholder* terkait untuk membuat keputusan yang lebih tepat (Azwanti, 2018; Pirmansyah & Wahyudi, 2023).

Dengan mengacu pada referensi-referensi tersebut, penelitian ini bertujuan untuk mengembangkan sistem berbasis algoritma C4.5 yang dapat mengolah data sensus ekonomi penduduk secara lebih detail dan akurat dalam menentukan kelayakan penerima bantuan sosial. Diharapkan hasil penelitian ini dapat memberikan kontribusi nyata dalam meningkatkan keadilan dan efektivitas distribusi dana bantuan sosial di Kelurahan Bangka Belitung.

2. Metode Penelitian

Jenis Penelitian, Waktu, dan Tempat

Penelitian ini merupakan penelitian kuantitatif dengan pendekatan deskriptif analitis yang bertujuan untuk mengembangkan model klasifikasi kelayakan penerima dana bantuan sosial menggunakan algoritma *Decision Tree*, khususnya C4.5. Penelitian ini dilakukan selama periode Januari hingga Juni 2024 di Kelurahan Bangka Belitung Laut, Kecamatan Pontianak Tenggara, Kalimantan Barat.

Target dan Subjek Penelitian

Target penelitian ini adalah seluruh penduduk Kelurahan Bangka Belitung Laut yang terdaftar dalam data sensus ekonomi tahun 2021. Subjek penelitian mencakup

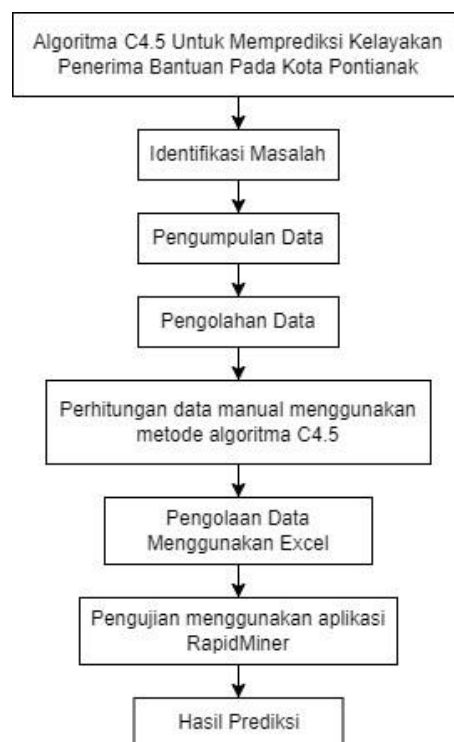
data penduduk yang terdiri dari berbagai variabel ekonomi dan sosial, termasuk jenis pekerjaan, pendapatan, dan status penerimaan bantuan sosial. Teknik *sampling* yang digunakan adalah *simple random sampling*, dimana dari populasi sebanyak 5167 penduduk, sampel yang diambil berjumlah 98 orang. Pemilihan sampel ini menggunakan rumus Slovin dengan tingkat kesalahan 10% (Septiani, Arribe, & Diansyah, 2020).

Prosedur Penelitian, Instrumen, dan Teknik Analisis Data

Prosedur penelitian dimulai dengan pengumpulan data melalui studi pustaka dan sensus ekonomi penduduk. Studi pustaka dilakukan untuk mengumpulkan data sekunder dari jurnal ilmiah, artikel, dan literatur terkait penggunaan algoritma *Decision Tree* dalam klasifikasi data. Sementara itu, data sensus ekonomi penduduk dikumpulkan dari Kelurahan Bangka Belitung Laut. Instrumen yang digunakan dalam penelitian ini termasuk kuesioner dengan skala *Likert* untuk mengukur variabel sosial dan ekonomi penduduk, serta aplikasi Rapid Miner untuk pengolahan data dan implementasi algoritma C4.5.

Setelah data terkumpul, dilakukan *preprocessing* data untuk membersihkan dan menyiapkan data sebelum dianalisis. Data kemudian diolah menggunakan algoritma C4.5 dengan aplikasi Rapid Miner, di mana model klasifikasi dihasilkan melalui metode *10-fold cross-validation* untuk memastikan akurasi yang tinggi. Teknik analisis data yang digunakan adalah analisis deskriptif untuk mengolah dan menginterpretasi data yang dikumpulkan. Hasil analisis disajikan dalam bentuk tabel, grafik, dan diagram untuk memudahkan interpretasi dan pengambilan keputusan. Model pohon keputusan yang dihasilkan

digunakan untuk menentukan kelayakan penerima bantuan berdasarkan variabel-variabel yang relevan seperti jenis pekerjaan dan pendapatan.



Gambar 1. Tahapan Penelitian
Sumber: Hasil Penelitian (2024)

3. Hasil dan Pembahasan

3.1. Hasil Penelitian

3.1.1. Analisis Data Penduduk

Data penduduk dianalisis berdasarkan berbagai kategori pekerjaan dan status penerimaan bantuan. Tabel 1 berikut menunjukkan distribusi penduduk berdasarkan jenis pekerjaan dan status penerimaan bantuan.

Tabel 1. Data Penduduk Kelurahan Bangka Belitung Laut Menurut Pekerjaan dan Status Penerimaan Bantuan

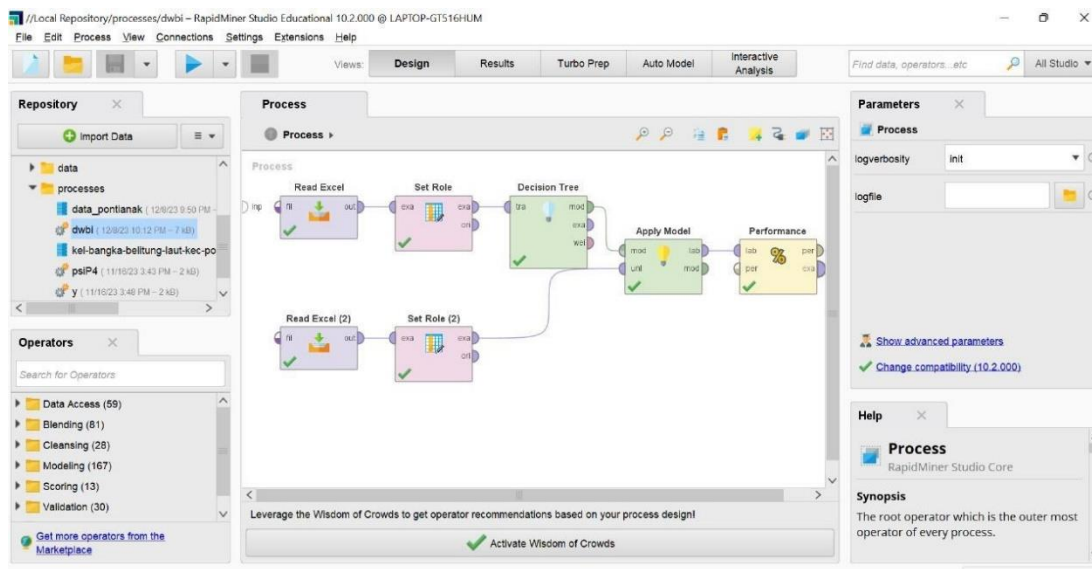
No	Jenis Pekerjaan	Jumlah	Menerima Bantuan	Tidak Menerima Bantuan
1	Belum/Tidak Bekerja	4.375	2.875	1.500
2	Aparatur Pejabat Negara	884	120	764
3	Tenaga Pengajar	129	45	84
4	Wiraswasta	3.873	1.400	2.473
5	Pertanian dan Peternakan	29	15	14
6	Nelayan	11	7	4
7	Pelajar dan Mahasiswa	2.242	1.120	1.122
8	Tenaga Kesehatan	30	5	25
9	Pensiunan	167	55	112
10	Pekerjaan Lainnya	2.563	1.200	1.363

Sumber: Data Sensus Ekonomi Kelurahan Bangka Belitung Laut (2021)

3.1.2. Pengolahan Data dengan Algoritma C4.5

Data yang telah dikumpulkan kemudian diolah menggunakan algoritma

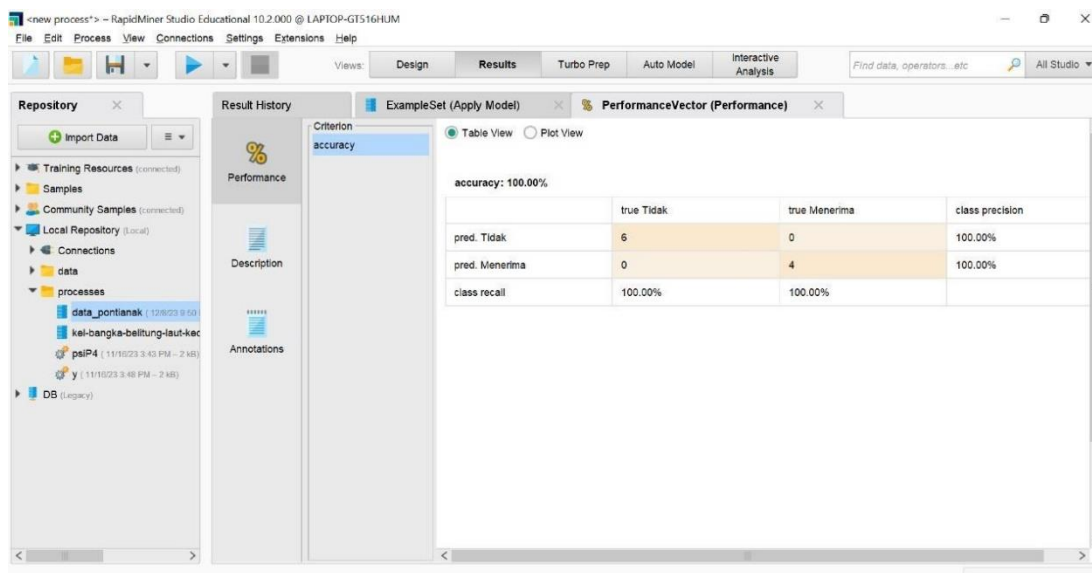
C4.5. Proses *preprocessing* data dilakukan untuk membersihkan dan menyiapkan data. Gambar berikut menunjukkan proses *preprocessing* data.



Gambar 2. *Preprocessing*
Sumber: Hasil Penelitian (2024)

Berdasarkan hasil olah data melalui aplikasi Rapid Miner, diperoleh *Confusion Matrix* untuk mengukur tingkat akurasi dari

algoritma C4.5. Hasil evaluasi menunjukkan bahwa algoritma ini mencapai akurasi, *precision*, dan *recall* sebesar 100%.

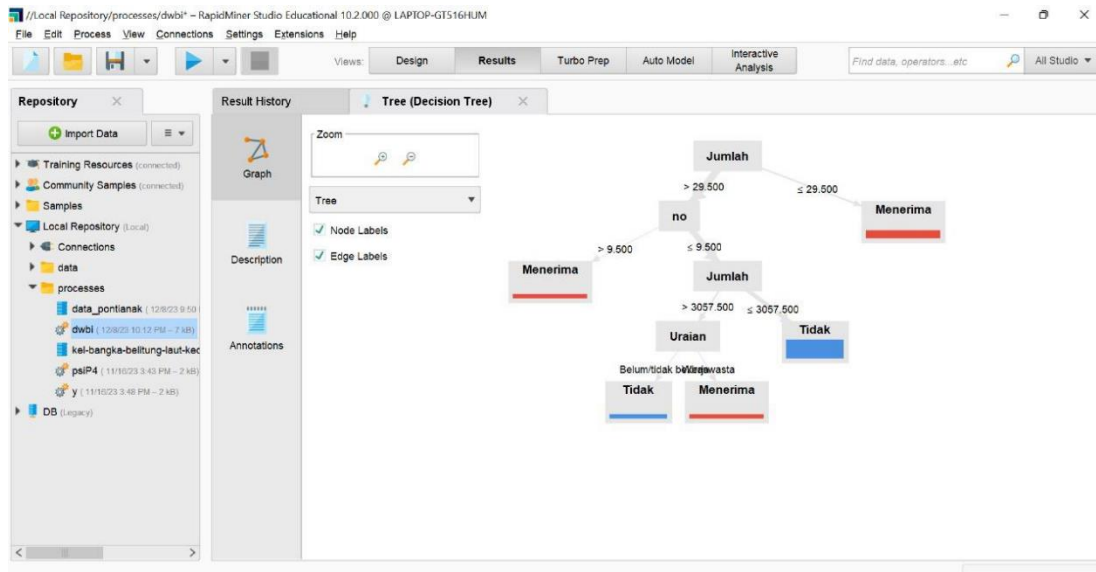


Gambar 3. *Accuracy*
Sumber: Hasil Penelitian (2024)

3.1.3. Model Pohon Keputusan

Model pohon keputusan yang dihasilkan dari algoritma C4.5 ditunjukkan pada Gambar 4 berikut. Model ini membantu

dalam menentukan kelayakan penerima bantuan berdasarkan variabel-variabel yang relevan seperti jenis pekerjaan dan pendapatan.



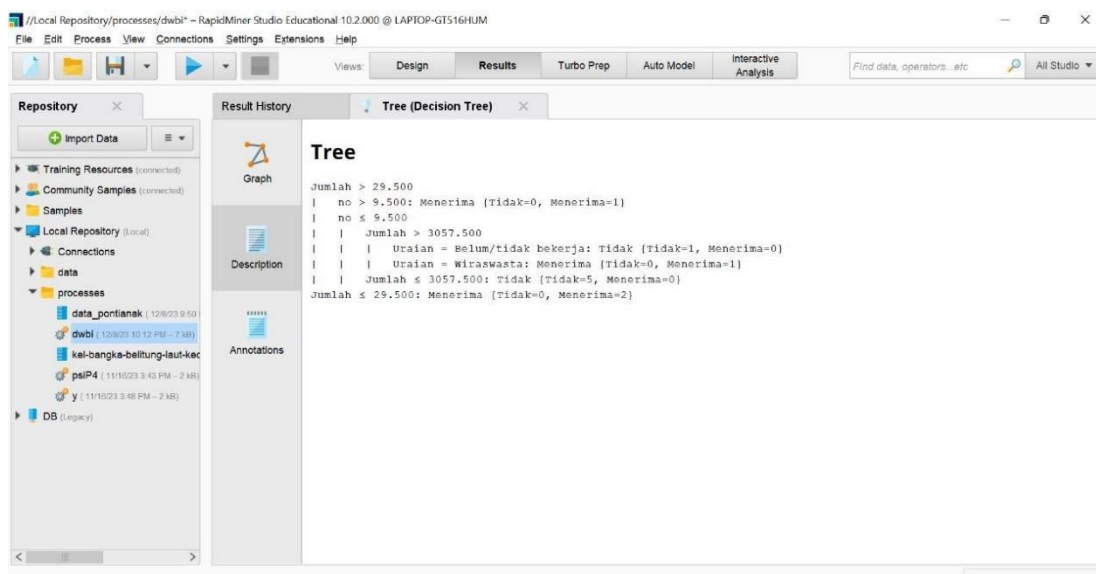
Gambar 4. Pohon Keputusan (Tree)
Sumber: Hasil Penelitian (2024)

Penjelasan dari model pohon keputusan yang dihasilkan: atribut utama yang menjadi akar adalah "Jenis Pekerjaan", diikuti dengan cabang-cabang yang menunjukkan pendapatan dan variabel lain yang mempengaruhi keputusan penerimaan bantuan.

3.1.4. Rule dari Pohon Keputusan

Berdasarkan model pohon keputusan, beberapa aturan (*rules*) yang dihasilkan adalah:

1. Jika jenis pekerjaan adalah "Belum/Tidak Bekerja" dan pendapatan $<$ Rp29.500, maka penerima layak menerima bantuan.
2. Jika jenis pekerjaan adalah "Wiraswasta" dan pendapatan $>$ Rp29.500, maka penerima tidak layak menerima bantuan.



Gambar 5. Description Tree
Sumber: Hasil Penelitian (2024)

3.2. Pembahasan

Hasil penelitian menunjukkan bahwa algoritma C4.5 memiliki kemampuan

yang sangat baik dalam mengklasifikasikan kelayakan penerima bantuan sosial. Dengan akurasi, *precision*, dan *recall* yang mencapai 100%, algoritma ini terbukti efektif dalam mengatasi masalah ketidaktepatan distribusi bantuan yang selama ini terjadi. Keakuratan model klasifikasi yang tinggi menunjukkan validitas model yang baik. Model ini mampu memprediksi kelayakan penerima bantuan dengan tepat, yang diindikasikan oleh tidak adanya kesalahan klasifikasi dalam *Confusion Matrix*. Algoritma C4.5 tidak hanya efektif tetapi juga mudah diinterpretasikan. Model pohon keputusan yang dihasilkan dapat digunakan oleh pengambil kebijakan untuk memahami dan mengaplikasikan hasil analisis dalam konteks distribusi bantuan sosial. Implementasi algoritma ini di Kelurahan Bangka Belitung Laut dapat meningkatkan keadilan dalam distribusi bantuan. Dengan mempertimbangkan variabel ekonomi dan sosial yang lebih detail, bantuan dapat disalurkan kepada yang benar-benar membutuhkan, sehingga mengurangi ketidakpuasan dan potensi konflik sosial di masyarakat.

4. Kesimpulan

Berdasarkan hasil penelitian, implementasi algoritma C4.5 dalam klasifikasi kelayakan penerima bantuan sosial menunjukkan hasil yang sangat baik dengan akurasi, *precision*, dan *recall* mencapai 100%. Hal ini menunjukkan bahwa algoritma ini efektif dalam menentukan penerima bantuan yang layak berdasarkan data sensus ekonomi. Model pohon keputusan yang dihasilkan dapat diinterpretasikan dengan mudah dan memberikan aturan (*rules*) yang jelas dalam menentukan kelayakan penerima bantuan. Atribut utama yang digunakan dalam model ini adalah jenis pekerjaan dan pendapatan, yang terbukti relevan dalam klasifikasi kelayakan. Validitas model yang tinggi menunjukkan bahwa algoritma C4.5 mampu menangani *dataset* dengan baik dan menghasilkan prediksi yang akurat. Tidak ada kesalahan klasifikasi dalam *Confusion Matrix*, yang menegaskan keandalan model ini.

Penelitian ini berhasil mengidentifikasi dan mengimplementasikan solusi yang lebih adil dan akurat untuk penentuan penerima bantuan sosial. Dengan menggunakan data sensus ekonomi yang lebih detail, model ini mampu

mengatasi ketidakakuratan yang terjadi pada sistem distribusi bantuan sebelumnya yang hanya berbasis rekomendasi personal. Berdasarkan temuan penelitian ini, beberapa rekomendasi untuk langkah selanjutnya adalah model ini dapat diadopsi dan diterapkan di wilayah lain dengan penyesuaian terhadap data lokal masing-masing. Hal ini dapat membantu pemerintah daerah dalam meningkatkan keadilan distribusi bantuan sosial secara lebih luas. Disarankan untuk menggunakan data yang lebih detail dan terkini untuk meningkatkan akurasi model klasifikasi. Pengumpulan data secara berkala dapat membantu dalam menyesuaikan model dengan perubahan kondisi sosial dan ekonomi. Penting untuk memberikan pelatihan dan sosialisasi kepada aparat desa dan pemerintah daerah mengenai penggunaan model ini. Pemahaman yang baik tentang interpretasi hasil model akan membantu dalam pengambilan keputusan yang lebih tepat. Mengembangkan sistem berbasis web yang mengintegrasikan algoritma C4.5 untuk memudahkan akses dan penggunaan oleh berbagai pihak yang berkepentingan. Sistem ini dapat menjadi alat bantu yang efektif dalam proses penentuan kelayakan penerima bantuan.

Referensi

- Abdulhafedh, A. (2022). Comparison between Common Statistical Modeling Techniques Used in Research, Including: Discriminant Analysis vs Logistic Regression, Ridge Regression vs LASSO, and Decision Tree vs Random Forest. *OALib*, 09(02), 1–19. <https://doi.org/10.4236/oalib.1108414>
- Aulia, N., Suarna, N., & Prihartono, W. (2024). Klasifikasi Penentuan Penerima Program Indonesia Pintar Di Krwilbidikcam Gregeed Menggunakan Algoritma C4.5. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(6), 3913–3919. <https://doi.org/10.36040/jati.v7i6.8294>
- Azwanti, N. (2018). Algoritma C4.5 Untuk Memprediksi Mahasiswa Yang Mengulang Mata Kuliah (Studi Kasus Di Amik Labuhan Batu). *Simetris: Jurnal Teknik Mesin, Elektro Dan Ilmu Komputer*, 9(1), 11–22. <https://doi.org/10.24176/simet.v9i1.1627>

- Ba'abbad, I., Althubiti, T., Alharbi, A., Alfarsi, K., & Rasheed, S. (2021). A Short Review of Classification Algorithms Accuracy for Data Prediction in Data Mining Applications. *Journal of Data Analysis and Information Processing*, 09(03), 162–174. <https://doi.org/10.4236/jdaip.2021.93011>
- Ernamia, E. M. A., & Herliana, A. (2022). Analisis Sentimen Kuliah Daring Dengan Algoritma Naïve Bayes, K-Nn Dan Decision Tree. *Jurnal Responsif: Riset Sains Dan Informatika*, 4(1), 70–80. <https://doi.org/10.51977/jti.v4i1.614>
- Ladianto, P. (2022). Implementasi Data Mining Untuk Bantuan Dana Bansos Dengan Menggunakan Algoritma C4.5 Di Pemerintahan Kabupaten Empat Lawang. *JUPITER (Jurnal Penelitian Ilmu Dan Teknik ...)*, (13), 287–297. Retrieved from <https://jurnal.polsri.ac.id/index.php/jupiter/article/view/5075%0Ahttps://jurnal.polsri.ac.id/index.php/jupiter/article/download/5075/2189>
- Nur, A., Rohim, A., Purnamasari, A. I., & Ali, I. (2024). Komparasi Efektifitas Algoritma C4.5 Dan Naïve Bayes Untuk Menentukan Kelayakan Penerima Manfaat Program Keluarga Harapan (Studi Kasus : Kecamatan Cicalengka Kabupaten Bandung). *Jurnal Mahasiswa Teknik Informatika*, 8(2), 2355–2362.
- Pirmansyah, F., & Wahyudi, T. (2023). Implementasi Data Mining Menggunakan Algoritma C4.5 Untuk Prediksi Evaluasi Anggota Satuan Pengamanan Studi Kasus Pt. Yimm Pulogadung. *Jurnal Indonesia : Manajemen Informatika Dan Komunikasi*, 4(3), 1566–1580. <https://doi.org/10.35870/jimik.v4i3.370>
- Septiani, Y., Arribe, E., & Diansyah, R. (2020). MENGGUNAKAN METODE SEVQUAL (Studi Kasus : Mahasiswa Universitas Abdurrah Pekanbaru). *Jurnal Teknologi Dan Open Source*, 3(1), 131–143.
- Sugiono, Taufik, A., & Faizal Amir, R. (2020). Penerapan Penerapan Teknik Pso Over Sampling Dan Adaboost J48 Untuk Memprediksi Cacat Software. *Jurnal Responsif: Riset Sains Dan Informatika*, 2(2), 198–203. <https://doi.org/10.51977/jti.v2i2.249>