

Penerapan Algoritma *Decision Tree* dalam Menentukan Kelayakan Penerima Bantuan Program Keluarga Harapan

DOI: <http://dx.doi.org/10.35889/progresif.v22i1.3494>

Creative Commons License 4.0 (CC BY –NC)



Eka Praja Wiyata Mandala^{1*}, Dewi Eka Putri²

Teknik Informatika, Universitas Putra Indonesia YPTK, Padang, Indonesia

*e-mail Corresponding Author: ekaprajawm@upiyptk.ac.id

Abstract

Determining the eligibility of beneficiaries for the Family Hope Program still faces targeting inaccuracies due to limited data analysis and subjective decision-making. This study aims to develop a Decision Tree-based classification model to objectively determine Family Hope Program eligibility using data-driven approaches. The dataset includes socio-economic attributes such as income, DTKS status, school-age children, pregnant or breastfeeding mothers, elderly members, disabilities, and occupation. The data were processed through cleaning, transformation, and splitting into training and testing sets. Experimental results show that the model achieved a highest accuracy of 96.00%, along with high precision and recall values. The resulting decision tree structure also provides interpretable rules to support decision-making. These findings demonstrate that the Decision Tree method is effective in improving the accuracy and transparency of Family Hope Program beneficiary selection.

Keywords: *Decision Tree; Family Hope Program; Classification; Data Mining; Social Assistance*

Abstrak

Penentuan kelayakan penerima Program Keluarga Harapan masih menghadapi permasalahan ketidaktepatan sasaran akibat keterbatasan analisis data dan subjektivitas dalam pengambilan keputusan. Penelitian ini bertujuan mengembangkan model klasifikasi berbasis *Decision Tree* untuk menentukan kelayakan penerima Program Keluarga Harapan secara objektif dan berbasis data. Dataset yang digunakan mencakup atribut sosial ekonomi seperti pendapatan, status DTKS, usia sekolah, ibu hamil atau menyusui, lansia, disabilitas, dan pekerjaan. Data diproses melalui tahap pembersihan, transformasi, dan pembagian data latih dan uji. Hasil pengujian menunjukkan bahwa model mencapai akurasi tertinggi sebesar 96,00%, dengan presisi dan recall yang tinggi. Struktur pohon keputusan juga menghasilkan aturan yang mudah diinterpretasikan sebagai dasar pengambilan keputusan. Hasil ini membuktikan bahwa *Decision Tree* efektif untuk meningkatkan ketepatan dan transparansi penyaluran bantuan Program Keluarga Harapan.

Kata kunci: *Decision Tree; Program Keluarga Harapan; Klasifikasi; Data Mining; Bantuan Sosial*

1. Pendahuluan

Kemiskinan merupakan permasalahan sosial yang bersifat multidimensional dan hingga saat ini masih menjadi tantangan utama dalam pembangunan nasional di Indonesia, karena tidak hanya berkaitan dengan rendahnya pendapatan, tetapi juga keterbatasan akses terhadap pendidikan, layanan kesehatan, dan kesejahteraan sosial [1]. Kondisi tersebut menimbulkan kesenjangan sosial yang berdampak pada rendahnya kualitas sumber daya manusia dan berlanjutnya siklus kemiskinan antar generasi [2]. Berbagai penelitian menunjukkan bahwa kemiskinan di Indonesia masih memerlukan intervensi kebijakan yang terarah dan berkelanjutan agar dapat ditangani secara efektif dan tepat sasaran [3]. Oleh karena itu, upaya penanggulangan

kemiskinan menjadi isu strategis yang memerlukan pendekatan kebijakan dan dukungan analisis berbasis data.

Sebagai upaya menanggulangi kemiskinan tersebut, pemerintah Indonesia mengimplementasikan sejumlah kebijakan perlindungan sosial, termasuk Program Keluarga Harapan (PKH), yang mulai diimplementasikan sejak tahun 2007 [4]. Melalui PKH, pemerintah memberikan bantuan bersyarat kepada keluarga kurang mampu guna mendukung pemenuhan kebutuhan dasar dan peningkatan kesejahteraan di bidang pendidikan, kesehatan, dan kesejahteraan sosial [5]. Program ini dirancang untuk mendorong investasi sumber daya manusia agar mampu memutus rantai kemiskinan dalam jangka panjang [6]. Namun demikian, implementasi PKH di berbagai daerah masih menghadapi sejumlah permasalahan, antara lain ketidaktepatan sasaran penerima bantuan, keterbatasan validitas Data Terpadu Kesejahteraan Sosial (DTKS), serta rendahnya pemahaman masyarakat terhadap mekanisme dan tujuan program [7], [8]. Permasalahan tersebut menyebabkan efektivitas program belum optimal dan menimbulkan potensi ketimpangan dalam penyaluran bantuan sosial, sehingga diperlukan mekanisme penentuan kelayakan yang lebih objektif dan berbasis data.

Berbagai studi empiris menunjukkan bahwa PKH memiliki kontribusi positif dalam meningkatkan kesejahteraan keluarga penerima manfaat, terutama pada aspek kesehatan dan pendapatan, meskipun pengaruhnya terhadap pendidikan dan penurunan kemiskinan jangka panjang masih menunjukkan hasil yang beragam [9], [10]. Penelitian Nella dkk. (2022) mengkaji klasifikasi penerima PKH di Desa Mirip menggunakan algoritma *Decision Tree C4.5* dengan atribut dominan berupa keberadaan anak sekolah dan menghasilkan akurasi sebesar 71,1% [11]. Rohim dkk. (2024) membandingkan algoritma *Decision Tree C4.5* dan *Naïve Bayes* dalam menentukan kelayakan penerima PKH di Kecamatan Cicalengka dengan variabel sosial ekonomi seperti pendapatan dan status keluarga, dan menemukan bahwa *Naïve Bayes* memiliki akurasi lebih tinggi dibandingkan *C4.5* [12]. Kariyamin dkk. (2025) membandingkan *KNN* dan *Decision Tree C4.5* pada data penerima PKH Kabupaten Wakatobi dengan atribut sosial ekonomi keluarga dan menunjukkan bahwa *C4.5* lebih stabil dan interpretatif dengan akurasi 96,22% [13]. Penelitian Alyandi dkk. (2025) juga menggunakan data PKH Wakatobi dengan 33 atribut sosial ekonomi dan membuktikan bahwa *Decision Tree C4.5* memiliki performa klasifikasi yang tinggi dibandingkan *KNN* dan *Naïve Bayes* [14]. Musdalifah dkk. (2025) menerapkan *C4.5* pada data PKH Kabupaten Manokwari dengan atribut jumlah tanggungan, anak sekolah, ibu hamil, lansia, disabilitas, jenis rumah, dan pekerjaan, dan memperoleh akurasi sebesar 72,20% [15]. Sementara itu, Jupriyanto dkk. (2025) membandingkan *Naïve Bayes* dan *Decision Tree C4.5* pada data calon penerima PKH Desa Tenjolaya dengan atribut penghasilan, pendidikan, pekerjaan, dan kondisi kesehatan, serta memperoleh akurasi sebesar 93% untuk *C4.5* [16]. Hidayati dkk. (2025) menggunakan *Decision Tree C4.5* pada data calon penerima bantuan sosial Kabupaten Indramayu dengan atribut penghasilan, usia, dan jumlah tanggungan, dan menghasilkan akurasi sebesar 93,22% [17]. Meskipun penelitian-penelitian tersebut menunjukkan bahwa algoritma klasifikasi mampu membantu penentuan kelayakan penerima bantuan sosial, masih ditemukan variasi kinerja model serta potensi bias dalam proses seleksi, khususnya dalam mengidentifikasi rumah tangga yang benar-benar layak menerima bantuan.

Berdasarkan kondisi tersebut, penelitian ini mengusulkan pendekatan klasifikasi berbasis data mining dengan menggunakan algoritma *Decision Tree* untuk menentukan kelayakan penerima Program Keluarga Harapan (PKH). Pendekatan ini dipandang relevan karena mampu mengolah data sosial ekonomi secara sistematis serta menghasilkan aturan keputusan yang bersifat interpretatif dan mudah dipahami oleh pengambil kebijakan. Dalam konteks analisis data, data mining digunakan untuk mengidentifikasi pola dan informasi penting yang terkandung dalam kumpulan data yang kompleks dengan memanfaatkan teknik statistik, matematika, kecerdasan buatan, dan pembelajaran mesin [18], [19]. Salah satu teknik utama dalam data mining adalah klasifikasi, yaitu proses pembentukan model yang memetakan data ke dalam kelas tertentu berdasarkan atribut yang dimilikinya [20], [21]. Dalam ranah machine learning, *Decision Tree* dikenal sebagai algoritma yang dapat digunakan untuk klasifikasi maupun prediksi dengan mengandalkan struktur pohon berlapis, di mana node merepresentasikan atribut, cabang menunjukkan kondisi, dan leaf node menggambarkan kelas hasil [22]. Kebaruan penelitian ini terletak pada penerapan model klasifikasi *Decision Tree* untuk mengintegrasikan variabel-variabel sosial ekonomi keluarga dalam membentuk aturan keputusan kelayakan penerima PKH yang transparan dan berbasis data, sehingga diharapkan mampu meningkatkan ketepatan sasaran dan efektivitas penyaluran bantuan sosial.

2. Metodologi

2.1 Algoritma Decision Tree

Algoritma *Decision Tree* merupakan metode klasifikasi yang membangun model berbentuk struktur pohon untuk memetakan data ke dalam kelas tertentu. Proses pembentukan pohon dilakukan dengan memilih atribut terbaik sebagai simpul (*node*) berdasarkan nilai pemisahan tertinggi terhadap kelas target. Pada penelitian ini digunakan algoritma *Decision Tree* C4.5, yang menentukan pemilihan atribut menggunakan konsep *entropy* dan *information gain* [23]. Nilai entropi digunakan untuk mengukur tingkat ketidakpastian data dalam suatu himpunan dan dirumuskan sebagai:

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

dengan p_i menyatakan proporsi data pada kelas ke- i . Selanjutnya, *information gain* digunakan untuk menentukan atribut terbaik dalam proses pemisahan data, yang dirumuskan sebagai:

$$Gain(S, A) = Entropy(S) - \sum_{v \in Values(A)} \frac{|S_v|}{|S|} * Entropy(S_v) \quad (2)$$

dengan S adalah himpunan data awal, A adalah atribut, dan S_v adalah subset data untuk setiap nilai atribut A . Atribut dengan nilai *information gain* tertinggi dipilih sebagai *node* pemisah hingga terbentuk struktur pohon keputusan yang menghasilkan kelas akhir berupa “Layak” atau “Tidak Layak”.

2.2 Data dan Variabel

Data yang digunakan dalam penelitian ini merupakan data calon penerima Program Keluarga Harapan (PKH) yang memuat informasi sosial ekonomi rumah tangga. Dataset terdiri dari sejumlah atribut yang digunakan sebagai variabel *input* (fitur) dan satu atribut sebagai variabel *output* (target). Variabel input meliputi: pendapatan, status DTKS, jumlah anak usia sekolah, status ibu hamil atau menyusui, jumlah lansia, status disabilitas, dan jenis pekerjaan. Variabel output adalah status kelayakan penerima PKH yang terdiri dari dua kelas, yaitu Layak dan Tidak Layak. Jumlah sampel data yang digunakan sebanyak 50 data Kepala Keluarga, yang kemudian dibagi menjadi dua bagian, yaitu data latih (*training set*) dan data uji (*testing set*). Pembagian data dilakukan dengan beberapa skenario proporsi, seperti 50%:50%, 60%:40%, 70%:30%, 80%:20%, dan 90%:10%, untuk melihat pengaruh komposisi data terhadap kinerja model.

2.3 Teknik Validasi Performa Algoritma Decision Tree

Validasi performa algoritma *Decision Tree* dilakukan untuk mengetahui tingkat ketepatan model dalam mengklasifikasikan data kelayakan penerima PKH. Mekanisme pengujian dilakukan dengan melatih model menggunakan data latih, kemudian menguji hasilnya menggunakan data uji yang tidak terlibat dalam proses pelatihan. Kinerja model diukur menggunakan beberapa metrik evaluasi, yaitu akurasi, presisi, *recall*, dan *F1-score*. Akurasi digunakan untuk mengukur proporsi prediksi yang benar terhadap seluruh data uji, presisi menunjukkan tingkat ketepatan model dalam memprediksi kelas layak, *recall* mengukur kemampuan model dalam mengidentifikasi seluruh data yang benar-benar layak, sedangkan *F1-score* merupakan nilai harmonisasi antara presisi dan *recall*. Selain itu, evaluasi juga dilakukan menggunakan *confusion matrix* untuk mengetahui jumlah prediksi benar dan salah pada masing-masing kelas. Proses pengujian ini bertujuan untuk memvalidasi efektivitas algoritma *Decision Tree* dalam mengklasifikasikan kelayakan penerima PKH secara objektif dan berbasis data, sedangkan hasil pengujian dan analisis performa model disajikan pada bagian hasil dan pembahasan.

3. Hasil dan Pembahasan

3.1 Implementasi Algoritma

Penelitian ini dimulai dengan mengumpulkan dataset Kepala Keluarga di Desa Taluak Kecamatan Lintau Buo Kabupaten Tanah Datar Sumatera Barat sebanyak 50 Kepala Keluarga yang berpeluang memperoleh bantuan Program Keluarga Harapan. Setiap data KK direpresentasikan oleh sejumlah fitur sosial ekonomi yang relevan dengan kriteria PKH, serta satu fitur target sebagai penentu kelas yang dapat dilihat pada Tabel 1.

Table 1. Data calon penerima bantuan PKH dengan label kelayakan

KK	DTKS	Penda- patan (Rp.)	Usia Sekolah (Orang)	Ibu Hamil/ Menyusui	Usia Lanjut (Orang)	Disabilitas	Pekerjaan	Menerima (Target)
KK01	Terdaftar	550000	2	Menyusui Susu Formula	Tidak Ada	Semi Permanen	Petani / Pekebun	Layak
KK02	Tidak Terdaftar	650000	>3	Tidak Hamil / Menyusui	2	Tidak Permanen	Jualan	Tidak
KK03	Terdaftar	600000	1	Menyusui ASI	2	Tidak Ada	Petani / Pekebun	Layak
KK04	Proses Daftar	1000000	Tidak Ada	Tidak Hamil / Menyusui	Tidak Ada	Tidak Ada	Petani / Pekebun	Layak
KK05	Tidak Terdaftar	1000000	2	Menyusui Susu Formula	1	Permanen	Petani / Pekebun	Tidak
KK06	Terdaftar	1000000	>3	Hamil	2	Tidak Permanen	Mengurus Rumah Tangga Jualan	Layak
KK07	Terdaftar	1100000	1	Menyusui ASI	1	Semi Permanen		Layak
KK08	Tidak Terdaftar	1000000	Tidak Ada	Tidak Hamil / Menyusui	Tidak Ada	Tidak Ada	Bengkel Las	Tidak
KK09	Proses Daftar	1900000	>3	Menyusui Susu Formula	2	Semi Permanen	Mubaligh	Layak
KK10	Terdaftar	750000	2	Tidak Hamil / Menyusui	3	Permanen	Petani / Pekebun	Layak
....								
KK50	Terdaftar	1200000	>3	Menyusui ASI	3	Tidak Ada	Mengurus Rumah Tangga	Layak

Fitur DTKS (Data Terpadu Kesejahteraan Sosial) menunjukkan status administrasi rumah tangga dalam sistem kesejahteraan sosial dan terbagi ke dalam kategori Terdaftar, Proses Daftar, dan Tidak Terdaftar. Status ini menjadi indikator awal kelayakan, karena rumah tangga yang terdaftar atau sedang diproses memiliki peluang lebih besar untuk menerima bantuan. Fitur Pendapatan merepresentasikan kondisi ekonomi rumah tangga dengan variasi dari pendapatan rendah hingga menengah, di mana pendapatan yang lebih rendah cenderung berkorelasi dengan kelayakan penerima bantuan. Fitur Usia Sekolah menggambarkan jumlah anggota keluarga usia sekolah, sedangkan fitur Ibu Hamil/Menyusui merepresentasikan kondisi kesehatan ibu dalam rumah tangga. Kedua fitur ini merupakan komponen utama PKH di bidang pendidikan dan kesehatan. Selain itu, fitur Usia Lanjut dan Disabilitas menunjukkan keberadaan kelompok rentan dalam rumah tangga yang menjadi prioritas penerima bantuan sosial. Fitur Pekerjaan mencerminkan sumber penghasilan utama rumah tangga yang didominasi sektor informal, sehingga berkaitan erat dengan tingkat kerentanan ekonomi.

Secara umum, keterkaitan antar fitur menunjukkan bahwa rumah tangga dengan pendapatan rendah, status DTKS terdaftar atau dalam proses, serta memiliki komponen PKH seperti anak usia sekolah, ibu hamil/menyusui, lansia, atau disabilitas, lebih cenderung dikategorikan layak menerima bantuan. Sebaliknya, rumah tangga tanpa komponen PKH dan tidak terdaftar dalam DTKS memiliki kecenderungan lebih rendah untuk menerima bantuan. Fitur target Menerima digunakan sebagai variabel dependen dengan dua kelas, yaitu Layak dan Tidak. Label ini dibentuk berdasarkan kombinasi status DTKS dan keberadaan komponen PKH dalam rumah tangga. Distribusi kelas target yang relatif seimbang menjadikan dataset ini sesuai untuk penerapan metode klasifikasi, khususnya algoritma *Decision Tree*, karena mampu menghasilkan aturan keputusan yang objektif dan mudah diinterpretasikan dalam konteks penyaluran bantuan PKH.

Langkah selanjutnya adalah menelaah dan membersihkan data. Pada tahap ini dilakukan analisis awal untuk memahami karakteristik data, meliputi pemeriksaan tipe data, distribusi nilai setiap atribut, serta identifikasi data tidak konsisten. Atribut Pendapatan berbentuk numerik dan menunjukkan jumlah penghasilan bulanan keluarga. Atribut Usia Sekolah dan Usia Lanjut bersifat

numerik diskrit yang menyatakan jumlah anggota keluarga pada kategori usia tertentu. Sementara itu, atribut DTKS, Ibu Hamil/Menyusui, Disabilitas, Pekerjaan, dan Menerima merupakan atribut kategorikal yang menggambarkan kondisi sosial ekonomi keluarga. Selanjutnya dilakukan pemeriksaan distribusi nilai pada setiap atribut. Misalnya, pada atribut DTKS ditemukan tiga kategori utama yaitu Terdaftar, Tidak Terdaftar, dan Proses Daftar. Pada atribut Ibu Hamil/Menyusui terdapat empat kategori yaitu Hamil, Menyusui ASI, Menyusui Susu Formula, dan Tidak Hamil/Menyusui. Analisis ini penting untuk memastikan bahwa seluruh nilai yang muncul sesuai dengan definisi atribut dan tidak terdapat nilai yang menyimpang. Data juga diperiksa terhadap nilai kosong, inkonsistensi penulisan, dan potensi duplikasi. Berdasarkan dataset yang digunakan, seluruh baris data memiliki nilai lengkap pada setiap atribut sehingga tidak ditemukan *missing value*. Namun, ditemukan variasi penulisan pada beberapa atribut kategorikal, misalnya penggunaan spasi yang berbeda pada nilai pekerjaan seperti Petani / Pekebun, Buruh Harian, dan Mengurus Rumah Tangga, serta penulisan kategori seperti Tidak Ada dan Tidak Permanen pada atribut disabilitas. Perbedaan ini berpotensi memengaruhi proses pembentukan model sehingga perlu diseragamkan.

Selanjutnya, dilakukan pembersihan dengan menyeragamkan kategori, memperbaiki nilai yang tidak valid, serta memastikan setiap record memiliki informasi lengkap sehingga dataset siap digunakan dalam proses pembentukan model klasifikasi. Berdasarkan hasil penelaahan, dilakukan pembersihan data untuk memastikan bahwa seluruh atribut berada dalam format yang konsisten dan dapat diproses oleh algoritma Decision Tree. Proses pembersihan meliputi standardisasi nilai kategori, validasi rentang nilai numerik, dan penyiapan data untuk proses pemodelan. Atribut numerik seperti pendapatan, usia_sekolah, dan usia_lanjut diperiksa untuk memastikan tidak terdapat nilai negatif atau tidak logis. Seluruh nilai pendapatan berada dalam rentang yang realistis (Rp500.000 – Rp2.100.000). Nilai usia_sekolah dan usia_lanjut juga valid karena menggunakan kategori diskrit seperti 1, 2, 3, >3, atau Tidak Ada, yang menggambarkan jumlah anggota keluarga dalam kelompok usia tertentu. Untuk keperluan pemodelan, kategori Tidak Ada pada atribut numerik diskrit dikonversi menjadi nilai 0 agar dapat diproses secara matematis oleh algoritma Decision Tree. Setiap baris data mewakili satu keluarga penerima PKH. Pemeriksaan duplikasi dilakukan dengan membandingkan seluruh atribut pada setiap record. Hasilnya menunjukkan tidak terdapat data duplikat, sehingga seluruh data dapat digunakan dalam proses pemodelan.

Langkah selanjutnya adalah membangun model *decision tree*. Langkah pertama dalam pembangunan model adalah menentukan atribut yang berperan sebagai variabel input (fitur) dan variabel target (label). Pada dataset PKH, tujuh atribut sosio-ekonomi digunakan sebagai input, sedangkan satu atribut digunakan sebagai output seperti yang ditampilkan pada Tabel 2.

Table 2. Penentuan variabel input dan target

Vaiabel	Atribut
Input (X)	DTKS, Pendapatan, Usia Sekolah, Ibu Hamil/Menyusui, Usia Lanjut, Disabilitas, Pekerjaan
Target (Y)	Menerima (Layak / Tidak Layak)

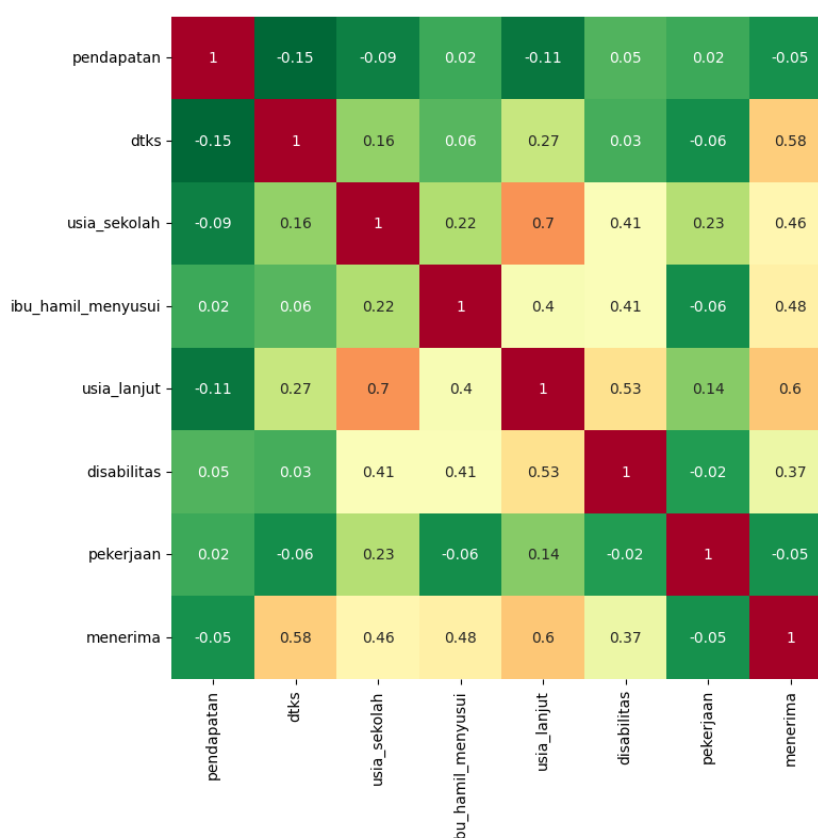
Dari Tabel 2 dapat dilihat bahwa atribut Menerima menjadi tujuan klasifikasi, yaitu menentukan apakah suatu keluarga layak atau tidak layak menerima bantuan PKH berdasarkan kondisi sosial dan ekonominya. Sebelum model dibangun, data hasil pembersihan ditransformasikan ke format yang dapat diproses oleh algoritma *Decision Tree*. Atribut numerik seperti Pendapatan, Usia Sekolah, dan Usia Lanjut digunakan secara langsung, sedangkan atribut kategorikal seperti DTKS, Ibu Hamil/Menyusui, Disabilitas, dan Pekerjaan dikonversi ke dalam bentuk numerik melalui teknik *encoding*. Transformasi ini bertujuan agar seluruh data dapat diproses secara matematis dalam pembentukan pohon keputusan. Hasil transformasi dapat dilihat pada Tabel 3.

Table 3. Hasil transformasi data calon penerima bantuan PKH

KK	DTKS	Penda- patan (Rp.)	Usia Sekolah (Orang)	Ibu Hamil/ Menyusui	Usia Lanjut (Orang)	Disabilitas	Pekerjaan	Menerima (Target)
KK01	2	550000	2	4	0	2	1	1
KK02	0	650000	4	1	2	1	2	0

KK	DTKS	Penda- patan (Rp.)	Usia Sekolah (Orang)	Ibu Hamil/ Menyusui	Usia Lanjut (Orang)	Disabilitas	Pekerjaan	Menerima (Target)
KK03	2	600000	1	3	2	0	1	1
KK04	1	1000000	0	1	0	0	1	1
KK05	0	1000000	2	4	1	3	1	0
KK06	2	1000000	4	2	2	1	3	1
KK07	2	1100000	1	3	1	2	2	1
KK08	0	1000000	0	1	0	0	4	0
KK09	1	1900000	4	4	2	2	5	1
KK10	2	750000	2	1	3	3	1	1
....								
KK50	2	1200000	4	3	3	0	3	1

Untuk memahami hubungan antar atribut dalam dataset penerima PKH, dilakukan analisis korelasi antar variabel. Analisis ini bertujuan untuk mengetahui seberapa kuat keterkaitan antara faktor sosial ekonomi keluarga dengan status kelayakan penerima bantuan. Hasil analisis korelasi divisualisasikan dalam bentuk heatmap sehingga pola hubungan antar atribut dapat diamati secara lebih jelas yang dapat dilihat pada Gambar 1.

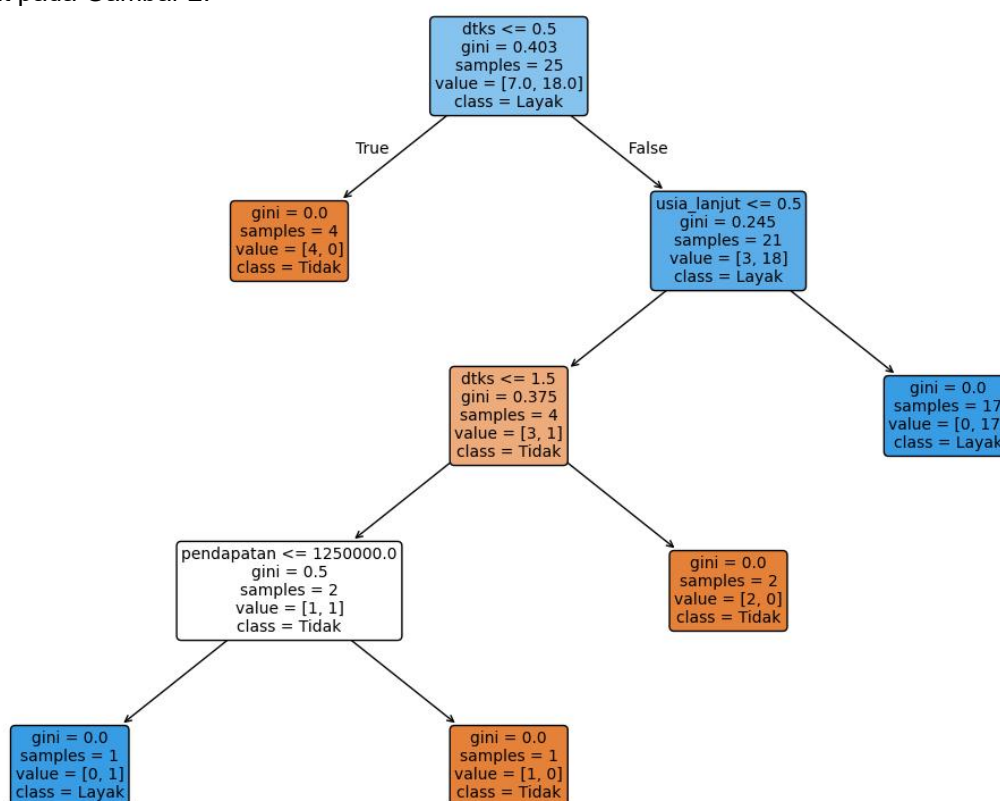


Gambar 1. Heatmap korelasi antar atribut dalam dataset penerima PKH

Gambar 1 menunjukkan nilai korelasi antara setiap pasangan atribut dalam dataset PKH, dengan rentang nilai dengan rentang nilai dari -1 sampai 1 , di mana nilai yang semakin mendekati 1 menunjukkan keterkaitan yang semakin kuat dan searah, sedangkan nilai yang mendekati 0 menandakan hubungan yang rendah. Terlihat bahwa atribut DTKS, Usia Lanjut, dan Usia Sekolah memiliki korelasi cukup tinggi terhadap variabel target menerima, masing-masing sebesar $0,58$, $0,60$, dan $0,46$. Hal ini menunjukkan bahwa status kepesertaan DTKS, jumlah lansia, dan jumlah anak usia sekolah merupakan faktor penting dalam menentukan kelayakan penerima bantuan PKH. Selain itu, atribut Ibu Hamil/Menyusui juga menunjukkan hubungan yang cukup kuat dengan nilai korelasi $0,48$, yang mengindikasikan bahwa kondisi ibu hamil atau

menyusui berpengaruh signifikan terhadap prioritas bantuan. Sebaliknya, atribut Pendapatan dan Pekerjaan memiliki korelasi yang sangat rendah terhadap variabel target, menandakan bahwa pada dataset ini, kedua faktor tersebut tidak menjadi pembeda utama dalam penentuan kelayakan. Secara keseluruhan, hasil korelasi ini menguatkan bahwa indikator kondisi keluarga dan status kepesertaan sosial lebih dominan dibandingkan aspek ekonomi semata dalam penentuan penerima Program Keluarga Harapan.

Pada tahap berikutnya algoritma *Decision Tree* mulai membangun model dengan cara memilih atribut yang paling informatif sebagai simpul akar (*root node*). Pemilihan atribut dilakukan berdasarkan nilai *information gain* atau *Gini index*, yang mengukur seberapa besar suatu atribut mampu membedakan kelas Layak dan Tidak Layak. Dalam konteks dataset PKH, atribut seperti Pendapatan, DTKS, dan Disabilitas cenderung memiliki nilai pemisahan yang tinggi karena sangat berpengaruh terhadap kelayakan penerima bantuan. Atribut dengan nilai terbaik dipilih sebagai simpul awal, kemudian data dibagi ke dalam cabang-cabang berdasarkan nilai atribut tersebut. Proses ini berlanjut secara rekursif hingga semua data pada satu cabang memiliki kelas yang sama atau tidak ada atribut yang dapat digunakan lagi. *Decision tree* yang dihasilkan dapat dilihat pada Gambar 2.



Gambar 2. Visualisasi Pohon Keputusan untuk Klasifikasi Kelayakan Penerima PKH

Gambar 2 menampilkan struktur pohon keputusan hasil pelatihan model *Decision Tree* untuk mengidentifikasi keluarga yang berhak menerima bantuan PKH. Akar pohon dimulai dari atribut DTKS, yang menunjukkan bahwa status kepesertaan dalam Data Terpadu Kesejahteraan Sosial merupakan faktor paling dominan dalam memisahkan data keluarga yang Layak dan Tidak Layak menerima bantuan. Jika nilai DTKS berada pada kategori tertentu (misalnya terdaftar), maka sebagian besar data langsung diklasifikasikan sebagai Layak, sedangkan pada kondisi lain keputusan akan diturunkan ke cabang berikutnya. Pada cabang selanjutnya, atribut Usia Lanjut digunakan untuk memisahkan keluarga berdasarkan keberadaan anggota lansia. Hal ini menunjukkan bahwa keberadaan lansia dalam keluarga menjadi indikator penting dalam penentuan kelayakan. Jika jumlah lansia memenuhi kriteria, maka keluarga cenderung diklasifikasikan sebagai Layak, sedangkan jika tidak, maka evaluasi dilanjutkan ke atribut lain seperti DTKS dan Pendapatan. Atribut Pendapatan muncul pada tingkat lebih bawah dalam pohon, yang berarti faktor ekonomi digunakan sebagai pemisah lanjutan ketika status sosial dan

kondisi keluarga belum cukup menentukan. Keluarga dengan pendapatan yang lebih rendah pada kondisi tertentu diklasifikasikan sebagai Layak, sementara yang berpendapatan lebih tinggi cenderung masuk kategori Tidak Layak. Aturan yang dihasilkan dari pohon keputusan dapat dilihat pada Tabel 4.

Table 4. Aturan klasifikasi kelayakan penerima PKH berdasarkan *Decision Tree*

No	Aturan	Interpretasi
1	IF DTKS = Tidak Terdaftar THEN Menerima = Tidak Layak	Keluarga yang tidak atau belum terdaftar DTKS langsung dikategorikan tidak layak
2	IF DTKS = (Terdaftar OR Proses Daftar) AND Usia Lanjut = (1 OR 2 OR 3 OR >3) THEN Menerima = Layak	Keluarga yang terdaftar/proses daftar DTKS dan memiliki lansia diprioritaskan menerima bantuan
3	IF DTKS = Proses Daftar AND Usia Lanjut = Tidak Ada AND Pendapatan ≤ Rp. 1.250.000 THEN Menerima = Layak	Keluarga dalam proses daftar DTKS, tanpa lansia, namun berpenghasilan rendah tetap layak
4	IF DTKS = Proses Daftar AND Usia Lanjut = Tidak Ada AND Pendapatan > Rp. 1.250.000 THEN Menerima = Layak	Keluarga dalam proses daftar DTKS, tanpa lansia, dan berpenghasilan relatif lebih tinggi tidak layak
5	IF DTKS = Terdaftar AND Usia Lanjut = Tidak Ada THEN Menerima = Tidak Layak	Keluarga dengan kategori DTKS terdaftar tanpa lansia tidak diprioritaskan menerima bantuan

Setiap simpul pada pohon menampilkan nilai gini, jumlah samples, dan distribusi kelas, yang menggambarkan tingkat kemurnian data pada setiap keputusan. Nilai gini yang semakin kecil pada simpul daun menunjukkan bahwa keputusan yang dihasilkan semakin pasti. Secara keseluruhan, pohon keputusan ini memberikan gambaran yang jelas dan mudah diinterpretasikan tentang bagaimana kombinasi faktor DTKS, kondisi keluarga, dan pendapatan digunakan oleh sistem untuk menentukan kelayakan penerima bantuan PKH secara objektif dan berbasis data.

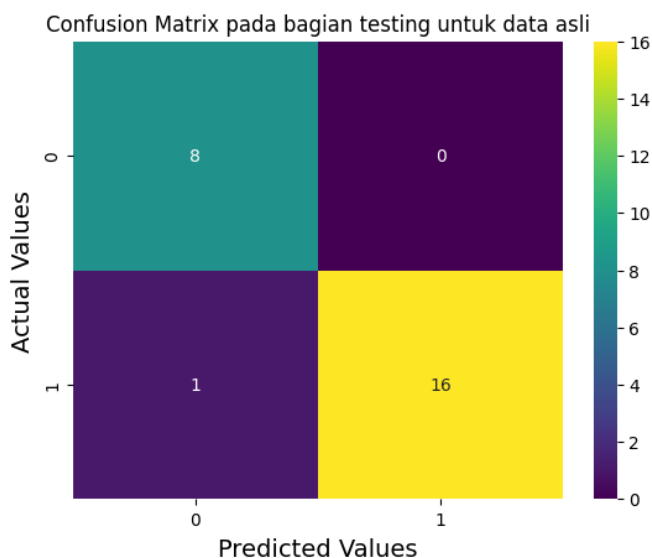
3.2 Pengukuran Validasi Efektifitas Kinerja Algoritma

Dataset selanjutnya dibagi menjadi dua bagian, yaitu data latih (*training set*) dan data uji (*testing set*). Data latih digunakan untuk membangun struktur pohon keputusan, sedangkan data uji digunakan untuk mengevaluasi kinerja model. Pembagian dilakukan dengan beberapa proporsi, mulai dari 50% untuk pelatihan dan 50% untuk pengujian sampai 90% untuk pelatihan dan 10% untuk pengujian sampai, agar model dapat diuji pada data yang belum pernah dilihat sebelumnya. Hasil pengujian dapat dilihat pada Tabel 5.

Table 5. Hasil pengujian pembagian data latih dan data uji

Splitting Data	Accuracy	Precision	Recall
50% : 50%	96.00 %	94.44 %	97.05 %
60% : 40%	95.00 %	92.86 %	96.43 %
70% : 30%	93.33 %	95.45 %	90.00 %
80% : 20%	90.00 %	87.50 %	92.86 %
90% : 10%	80.00 %	75.00 %	87.50 %

Tabel 5 menunjukkan hasil pengujian berbagai skenario pembagian data latih dan data uji terhadap kinerja model *Decision Tree*. Terlihat bahwa pembagian 50% data latih dan 50% data uji menghasilkan performa terbaik dengan nilai akurasi 96,00%, presisi 94,44%, dan recall 97,05%. Hasil ini menunjukkan bahwa model mampu mengenali keluarga yang layak menerima bantuan dengan sangat baik sekaligus menjaga tingkat kesalahan klasifikasi yang rendah. Seiring bertambahnya proporsi data latih, kinerja model justru mengalami penurunan, terutama pada skenario 80:20 dan 90:10. Hal ini mengindikasikan bahwa pada dataset yang relatif kecil, pembagian data yang terlalu tidak seimbang dapat menyebabkan model kehilangan kemampuan generalisasi. Oleh karena itu, skenario 50:50 dipilih sebagai konfigurasi terbaik untuk evaluasi lanjutan. Setelah diperoleh konfigurasi pembagian data terbaik, langkah selanjutnya adalah mengevaluasi performa model secara lebih rinci melalui *confusion matrix* pada data pengujian. Analisis ini digunakan untuk melihat secara langsung kemampuan model dalam mengklasifikasikan data layak dan tidak layak secara benar maupun salah. *Confusion matrix* data uji dapat dilihat pada Gambar 3.



Gambar 3. *Confusion matrix* model Decision Tree pada data uji penerima PKH

Gambar 3 menampilkan *confusion matrix* dari model *Decision Tree* pada data pengujian. Terlihat bahwa dari seluruh data uji, terdapat 8 data Tidak Layak yang berhasil diklasifikasikan dengan benar dan 16 data Layak yang juga diprediksi secara tepat oleh model. Hanya terdapat 1 data Layak yang salah diklasifikasikan sebagai Tidak Layak, sementara tidak ada data Tidak Layak yang salah diprediksi sebagai Layak. Hasil ini menunjukkan bahwa model memiliki kemampuan yang sangat baik dalam meminimalkan kesalahan pemberian bantuan kepada keluarga yang tidak memenuhi kriteria. Tingginya jumlah prediksi benar pada kelas Layak juga mengindikasikan bahwa model sangat efektif dalam mengidentifikasi keluarga yang berhak menerima Program Keluarga Harapan, sehingga sistem yang dibangun dapat menjadi alat bantu yang andal dalam proses penyaluran bantuan sosial secara tepat sasaran.

3.3 Pembahasan

Penelitian ini berhasil membangun model *Decision Tree* untuk mengklasifikasikan kelayakan penerima bantuan Program Keluarga Harapan (PKH) berdasarkan atribut sosio-ekonomi keluarga (misalnya DTKS, usia lanjut, usia sekolah, ibu hamil/menyusui, disabilitas, pendapatan, dan pekerjaan). Keberhasilan penelitian ditunjukkan oleh performa evaluasi yang tinggi pada skenario pembagian data terbaik, yakni akurasi 96,00%, presisi 94,44%, dan recall 97,05%. Selain itu, *confusion matrix* menunjukkan kesalahan klasifikasi yang sangat rendah, di mana prediksi model lebih dominan berada pada kategori benar, serta tidak muncul kesalahan yang berisiko “meloloskan” keluarga yang tidak layak menjadi layak. Pola ini mengindikasikan bahwa model efektif sebagai alat bantu seleksi penerima bantuan karena tidak hanya akurat, tetapi juga menjaga ketepatan sasaran program.

Dari sisi interpretabilitas, model *Decision Tree* menghasilkan aturan (*rules*) yang transparan. Pada pohon keputusan, atribut DTKS muncul sebagai pemisah awal (*root*), diikuti atribut seperti Usia Lanjut dan Pendapatan pada cabang tertentu. Ini memperkuat temuan analisis korelasi sebelumnya, bahwa variabel sosial-keluarga (DTKS dan keberadaan lansia) lebih dominan dalam membedakan kelayakan dibanding variabel lain pada dataset ini. Keunggulan interpretabilitas ini penting dalam konteks bantuan sosial karena pihak pengambil kebijakan dapat menelusuri alasan keputusan model secara jelas.

Jika dibandingkan dengan penelitian sejenis, akurasi model pada penelitian ini tergolong lebih tinggi. Penelitian sebelumnya menggunakan *Decision Tree* untuk klasifikasi status kemiskinan melaporkan akurasi sebesar 93%, dimana variabel pendapatan dan status pekerjaan menjadi penentu utama klasifikasi. Hasil ini menunjukkan bahwa meskipun *Decision Tree* mampu bekerja baik pada kasus klasifikasi sosial ekonomi, performanya sedikit lebih rendah dibandingkan hasil penelitian ini yang fokus pada kelayakan PKH karena desain fitur dan data yang berbeda [24]. Selain itu, penelitian lain yang menerapkan algoritma C4.5 pada dataset penerima bantuan PKH di Kecamatan Atur Lintang menghasilkan akurasi sekitar 94%, dengan

precision mencapai 0,96 dan recall 0,90. Meskipun metode tersebut juga menghasilkan model yang dapat digunakan secara efektif, kinerjanya secara statistik tetap berada di bawah akurasi 96% yang diperoleh dalam penelitian ini, sehingga semakin menegaskan efektivitas fitur dan preprocessing yang dilakukan pada dataset PKH yang digunakan [25]. Dibandingkan studi tersebut, model Decision Tree pada penelitian ini menunjukkan akurasi lebih tinggi (96%) dan recall lebih tinggi (97,05%) pada konfigurasi terbaik, yang berarti model ini lebih sensitif dalam menangkap keluarga yang benar-benar layak (mengurangi risiko “layak tapi terlewat”).

Secara metodologis, capaian akurasi 96,00% pada penelitian ini dapat dijelaskan oleh beberapa hal: (1) atribut yang digunakan relevan terhadap kebijakan PKH (misalnya DTKS dan indikator kondisi keluarga), (2) data telah melalui pembersihan dan standardisasi kategori, sehingga mengurangi noise pada fitur kategorikal, serta (3) pemilihan skenario pembagian data yang paling sesuai untuk karakteristik dataset menghasilkan generalisasi yang lebih stabil. Dibandingkan penelitian pembandingan yang akurasinya berada pada rentang 93–94%, model pada penelitian ini menunjukkan kemampuan klasifikasi yang lebih kuat pada dataset yang digunakan.

Meskipun hasil penelitian menunjukkan kinerja model yang tinggi, penelitian ini masih memiliki beberapa keterbatasan. Pertama, ukuran dataset yang digunakan relatif terbatas dan hanya merepresentasikan kondisi pada satu wilayah tertentu, sehingga kemampuan generalisasi model pada wilayah lain dengan karakteristik sosial ekonomi yang berbeda belum dapat dipastikan. Kedua, atribut yang digunakan masih terbatas pada variabel sosial ekonomi dasar, sehingga belum mencakup faktor lain yang berpotensi memengaruhi kelayakan penerima PKH, seperti kondisi geografis, akses terhadap layanan publik, maupun dinamika perubahan data dari waktu ke waktu. Ketiga, validasi model dilakukan dengan skema pembagian data latih dan data uji, sehingga belum sepenuhnya menggambarkan stabilitas model pada variasi data yang lebih luas.

Berdasarkan keterbatasan tersebut, penelitian selanjutnya direkomendasikan untuk menggunakan dataset dengan cakupan wilayah dan jumlah sampel yang lebih besar, serta menambahkan atribut yang lebih beragam agar model mampu merepresentasikan kondisi sosial ekonomi secara lebih komprehensif. Pengembangan lebih lanjut juga dapat diarahkan pada integrasi model ke dalam sistem pendukung keputusan berbasis aplikasi, sehingga hasil klasifikasi dapat dimanfaatkan secara langsung dalam proses penentuan penerima bantuan PKH secara berkelanjutan.

4. Simpulan

Penelitian ini menyimpulkan bahwa algoritma *Decision Tree* mampu digunakan secara efektif untuk mengklasifikasikan kelayakan penerima Program Keluarga Harapan (PKH) berdasarkan data sosial ekonomi keluarga. Melalui tahapan pengumpulan data, pembersihan, pemodelan, dan evaluasi, diperoleh model dengan kinerja yang sangat baik, ditunjukkan oleh akurasi mencapai 96,00%, serta nilai presisi mencapai 94,44% dan recall yang tinggi yaitu 97,05%. Hasil ini membuktikan bahwa atribut seperti status DTKS, keberadaan lansia, kondisi ibu hamil atau menyusui, dan tingkat pendapatan memiliki peran penting dalam menentukan kelayakan bantuan, dan hubungan antar atribut tersebut dapat dipetakan secara jelas melalui struktur pohon keputusan yang mudah diinterpretasikan. Dengan demikian, penelitian ini tidak hanya menjawab tujuan untuk membangun model klasifikasi kelayakan PKH yang akurat, tetapi juga menghasilkan kerangka keputusan berbasis aturan yang transparan dan dapat dijadikan dasar pertimbangan bagi pihak pengambil kebijakan dalam penyaluran bantuan sosial. Untuk penelitian selanjutnya, model ini memiliki prospek untuk dikembangkan lebih lanjut dengan penambahan jumlah dan ragam data, penerapan teknik validasi silang, serta integrasi dengan sistem informasi kesejahteraan sosial agar proses seleksi penerima bantuan dapat dilakukan secara lebih adaptif, objektif, dan berkelanjutan.

Daftar Referensi

- [1] W. S. Wahyuni dan F. P. Putra, “Analisis Karakteristik Rumah Tangga Miskin Penerima Program Keluarga Harapan dan Penerima Keluarga Harapan Sekaligus Bantuan Pangan Non-Tunai di Indonesia,” *J. Inform. Ekon. Bisnis*, vol. 7, no. 2, pp. 141–147, 2025, doi: 10.37034/infekon.v7i2.1112.
- [2] D. M. Yunisar, I. Murti, dan S. Hartono, “Evaluasi Program Keluarga Harapan Dalam Penurunan Kemiskinan Kelurahan Sidotopo Kota Surabaya,” *PRAJA Obs. J. Penelit. Adm.*

- Publik*, vol. 5, no. 04, pp. 59–68, 2025, [Daring]. Tersedia pada: <https://aksiologi.org/index.php/praja/article/view/2291>.
- [3] A. K. D. Permata dan R. Parianom, "Analisis Pendapatan, Pendidikan, Kesehatan Program Keluarga Harapan Terhadap Kesejahteraan Masyarakat Miskin Di Jawa Barat," *J. Dev. Econ. Digit.*, vol. 4, no. 1, pp. 77–97, 2025, [Daring]. Tersedia pada: <https://ejournal.upnvj.ac.id/jded/article/view/8974>.
 - [4] R. Mahmudah, A. Baihaqi, dan Jumaidi, "Efektivitas Program Keluarga Harapan (PKH) Dalam Mengurangi Tingkat Kemiskinan Di Desa Mihu Kecamatan Juai Kabupaten Balangan," *J. Kebijak. Publik*, vol. 1, no. 2025, pp. 411–422, 2M, [Daring]. Tersedia pada: <https://ejurnal.stiaamuntai.ac.id/index.php/PPJ/article/view/993>.
 - [5] E. T. Andi, A. K. Andi, N. K. Hanafie, dan Najamuddin, "Efektivitas Program Keluarga Harapan Dalam Perspektif Permensos Nomor 1 Tahun 2018 Bagi Keluarga Penerima Manfaat," *J. Educ. Dev.*, vol. 13, no. 2, pp. 494–499, 2025, doi: 10.37081/ed.v13i2.6827.
 - [6] E. Abiyani dan E. Setyowati, "Efektivitas Bantuan Sosial Program Keluarga Harapan (PKH) di Kabupaten Magetan berdasar Prinsip Keadilan dalam Islam," *Stud. Econ. J. Ekon. Islam*, vol. 11, no. 2, pp. 168–180, 2025, [Daring]. Tersedia pada: <https://jurnal.uinsu.ac.id/index.php/studiaeconomica/article/view/25806>.
 - [7] F. Angraini, N. Hasanah, dan E. Pratama, "Implementasi Kebijakan Program Keluarga Harapan di Kota Dumai," *J. Adm. Publik Bisnis*, vol. 7, no. 2, pp. 323–334, 2025, doi: 10.36917/japabis.v7i2.389.
 - [8] R. Hanavi dan S. M. Sihombing, "Implementasi Program Keluarga Harapan (PKH) dalam Upaya Meningkatkan Kesejahteraan Sosial Ekonomi Keluarga Miskin di Kelurahan Pekan Binjai Kecamatan Binjai Kota Binjai (Studi Kasus pada Dinas Sosial Kota Binjai dan Kantor Kelurahan Pekan Binjai)," *Al-Faqih J. Ilmu Sos. Humaniora, Tek.*, vol. 1, no. 4, pp. 74–85, 2025, [Daring]. Tersedia pada: <https://journal.salahuddinal-ayyubi.com/index.php/AFJISH/article/view/544>.
 - [9] A. Ahmad, M. A. Arham, dan R. F. Djuuna, "Analisis Pengaruh Kebijakan Program Keluarga Harapan (PKH) Terhadap Kemiskinan Di Provinsi Gorontalo," *J. Akuntansi, Manaj. dan Ilmu Ekon.*, vol. 05, no. 03, pp. 398–410, 2025, doi: 10.54209/jasmien.v5i03.1265 Hal.
 - [10] R. M. Siregar, I. Puansah, dan N. Parapat, "Efektivitas Program Keluarga Harapan Dalam Meningkatkan Kesejahteraan Keluarga Miskin Di Desa Sibatuloting," *J. Public Issues J. Ilmu Sos. Humaniora, Polit. Dan Kebijakan Publik*, vol. 2, no. 3, pp. 1–5, 2025, [Daring]. Tersedia pada: <https://jurnal.puansahmediagroup.id/index.php/publicissues/article/view/19>.
 - [11] N. I. Nella, N. Y. Setiawan, dan D. E. Ratnawati, "Klasifikasi Penerima Bantuan Program Keluarga Harapan menggunakan Algoritme Decision Tree C4.5 (Studi Kasus : Desa Mlirip Kabupaten Mojokerto)," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 6, no. 3, pp. 1332–1339, 2022, [Daring]. Tersedia pada: <https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/10801>.
 - [12] A. N. A. Rohim, A. I. Purnamasari, dan I. Ali, "Komparasi Efektifitas Algoritma C4.5 Dan Naïve Bayes Untuk Menentukan Kelayakan Penerima Manfaat Program Keluarga Harapan (Studi Kasus : Kecamatan Cicalengka Kabupaten Bandung)," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 8, no. 2, pp. 2355–2362, 2024, doi: 10.36040/jati.v8i2.8345.
 - [13] Kariyamin, L. O. Alyandi, dan S. Arif, "Perbandingan Algoritma K-Nearest Neighbour dan C4.5 Decision Tree Untuk Klasifikasi Penerima Bantuan Program Keluarga Harapan," *Decod. J. Pendidik. Teknol. Inf.*, vol. 5, no. 3, pp. 982–992, 2025, doi: 10.51454/decode.v5i3.1353.
 - [14] L. O. Alyandi, Kariyamin, dan S. Arif, "Perbandingan Algoritma KNN, C4.5 Dan Naïve Bayes Untuk Klasifikasi Penerima Bantuan Program Keluarga Harapan," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 9, no. 6, pp. 9227–9234, 2025, doi: 10.36040/jati.v9i6.15531.
 - [15] M. Daud, R. Juita, dan C. D. Suhendra, "Penerapan Metode Algoritma C4.5 Untuk Klasifikasi Kelayakan Penerima Program Bantuan Pada Dinas Sosial Kabupaten Manokwari," *Decod. J. Pendidik. Teknol. Inf.*, vol. 5, no. 1, pp. 271–278, 2025, doi: 10.51454/decode.v5i1.1057.
 - [16] Jupriyanto dkk., "Perbandingan Kinerja Algoritma Naïve Bayes Dan C4.5 Pada Sistem Web Klasifikasi Kelayakan PKH," *J. Teknol. Inf. dan Komun.*, vol. 18, no. 1, pp. 50–64, 2025, doi: 10.47561/jtik.v18i1.287.
 - [17] N. Hidayati, N. Suarna, I. Ali, dan D. Solihudin, "Implementasi Algoritma C4.5 untuk Meningkatkan Akurasi Klasifikasi Penerima Bantuan Sosial di Indramayu," *J. Comput. Sci. Artif. Intell.*, vol. 02, no. 1, pp. 1–8, 2025, [Daring]. Tersedia pada:

- <https://ruangjurnal.or.id/index.php/jcsai/article/view/6>.
- [18] D. W. Safitri, E. Fadilah, dan S. Rahayu, "Klasifikasi Kelayakan Penerima Bantuan Program Keluarga Harapan Di Kecamatan Sinar Peninjauan Menggunakan Metode Naïve Bayes," *Simtek J. Sist. Inf. Dan Tek. Komput.*, vol. 10, no. 2, pp. 361–368, 2025, doi: 10.51876/simtek.v10i2.1631.
- [19] D. C. Alhafidz, M. Hatta, dan M. Febima, "Perancangan Sistem Prediksi Bantuan PKH di Dinsos Cirebon Menggunakan Naive Bayes," *J. Artif. Intell. Digit. Bus.*, vol. 4, no. 3, pp. 3633–3646, 2025, [Daring]. Tersedia pada: <https://journal.ilmudata.co.id/index.php/RIGGS/article/view/2526>.
- [20] Mursyida, G. N. A. Wibawa, dan Arman, "Penerapan Metode Naive Bayes Untuk Identifikasi Penerima Bantuan Program Keluarga Harapan (PKH) Di Kabupaten Buton Tengah," *J. Mat. Komputasi dan Stat.*, vol. 5, no. 2, pp. 1003–1009, 2025, [Daring]. Tersedia pada: <https://jmks.uho.ac.id/index.php/journal/article/view/147>.
- [21] H. N. F. Fikrillah, "Klasifikasi Program Bantuan Sosial Menggunakan Algoritma K- Nearest Neighbor (K-NN) (Studi Kasus Kecamatan Malangbong Kabupaten Garut)," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 11, no. 3, pp. 609–616, 2025, doi: 10.25126/jtiik.2025128817.
- [22] A. I. Perhan, I. Yustiana, dan I. Sanjaya, "Implementation of Machine Learning Using Decision Tree Method for Social Assistance Recipient Classification," *bit-Tech*, vol. 8, no. 1, pp. 901–909, 2025, doi: 10.32877/bt.v8i1.2755.
- [23] F. Rezkika, B.N. Sari, & A.S.Y. Irawan, "Klasifikasi Masa Tunggu Alumni Untuk Mendapatkan Pekerjaan Berdasarkan Kompetensi Menggunakan Algoritma C4. 5 (Studi Kasus: Fasilkom Unsika). *Progresif: Jurnal Ilmiah Komputer*, vol. 17, no. 2, pp. 95-106, 2021.
- [24] A. Setyawan, A. Fitriani, dan E. Rilvani, "Klasifikasi Kemiskinan di Indonesia dengan Decision Tree Menggunakan Rapidminer," *J. Media Akad.*, vol. 3, no. 7, pp. 1–12, 2025, doi: 10.62281/v3i7.2646.
- [25] S. Wahyuni, Asrianda, dan S. Retno, "Classification of Family Hope Program Assistance Recipients Using the C4.5 Algorithm with Z-Score Normalization (Case Study in Atu Lintang District)," *Inf. Technol. Eng. Journals*, vol. 10, no. 1, pp. 160–173, 2025, doi: 10.24235/itej.v10i1.207.