

KLASIFIKASI PENENTUAN PENERIMA PROGRAM INDONESIA PINTAR DI KRWILBIDIKCAM GREGED MENGGUNAKAN ALGORITMA C4.5

Nurul Aulia, Nana Suarna, Willy Prihartono

Teknik Informatika, STMIK IKMI Cirebon
Jl Perjuangan No. 10B Karyamulya Cirebon, Indonesia
nurulauliamei10@gmail.com

ABSTRAK

Program Indonesia Pintar merupakan program dukungan pendidikan yang bertujuan untuk meningkatkan akses dan mutu pendidikan bagi siswa sekolah dasar. Penelitian ini bertujuan untuk mengimplementasikan algoritma C4. 5 dalam rangka mengidentifikasi penerima manfaat Program Indonesia Pintar (PIP) di Koorwilbidikcam GregeDUntuk mengoptimalkan penyaluran dana PIP, diperlukan klasifikasi berdasarkan data mining untuk memahami dan mengidentifikasi faktor-faktor kunci yang mempengaruhi identifikasi penerima manfaat program ini. Metode penelitian ini meliputi langkah pengumpulan data, preprocessing data, pelatihan model menggunakan algoritma C4. 5, evaluasi model, dan interpretasi hasil. Dengan menggabungkan teknik data mining dan analisis kriteria untuk mengidentifikasi penerima manfaat PIP. Dengan menggabungkan teknik data mining dan analisis kriteria untuk mengidentifikasi penerima manfaat PIP, penelitian ini diharapkan dapat berkontribusi dalam meningkatkan efektivitas dan efisiensi pelaksanaan program dukungan, tunjangan pengangguran, dan pendidikan di tingkat dasar. Hasil penelitian ini diharapkan dapat memberikan pemahaman yang lebih baik mengenai faktor-faktor yang mempengaruhi identifikasi penerima manfaat PIP di wilayah Koorwilbidikcam GregeD dan memberikan landasan yang kokoh dalam pengambilan keputusan yang lebih cerdas dalam menyalurkan dana PIP.

Kata kunci: Klasifikasi, Program Indonesia Pintar, Algoritma C4.5, variable, interpretasi, data mining

1. PENDAHULUAN

Pendidikan memegang peranan penting dalam pembangunan suatu negara. Program Indonesia Pintar (PIP) merupakan salah satu inisiatif pemerintah untuk meningkatkan akses dan mutu pendidikan khususnya pada tingkat dasar (SD). Program Indonesia Pintar (PIP) merupakan dukungan pemerintah dalam bentuk pembayaran langsung kepada siswa sesuai kriteria yang telah ditentukan. Tujuan PIP adalah memberikan akses pendidikan dan pendidikan bermutu serta menurunkan angka putus sekolah. PIP memberikan dukungan finansial kepada siswa dengan tujuan mengurangi beban biaya pendidikan dan mendorong siswa untuk mengikuti pendidikan formal. Beasiswa adalah pemberian berupa bantuan keuangan yang diberikan pada perorangan yang bertujuan untuk dipergunakan demi keberlangsungan pendidikan yang ditempuh

Penelitian ini bertujuan untuk meningkatkan ketepatan dan akurasi dalam penentuan penerima manfaat Program Indonesia Pintar (PIP) di Koorwilbidikcam GregeD melalui implementasi algoritma C4.5. PIP merupakan inisiatif pemerintah untuk meningkatkan akses dan kualitas pendidikan, khususnya pada tingkat dasar, dengan memberikan dukungan finansial kepada siswa sesuai kriteria tertentu. Meskipun PIP memiliki dampak positif, ketidakpastian dalam mengidentifikasi penerima manfaat dapat menghambat efektivitas penyaluran bantuan. Kompleksitas kondisi ekonomi dan sosial siswa di Koorwilbidikcam GregeD memperumit proses identifikasi, sehingga diperlukan pendekatan cerdas

seperti algoritma C4.5. Penelitian ini merinci hasil penelitian terdahulu yang menggunakan algoritma C4.5 dalam konteks serupa. Dalam sebuah studi, algoritma ini berhasil menghasilkan aturan keputusan yang mampu meningkatkan ketepatan identifikasi penerima manfaat PIP. Namun, penelitian lain juga menunjukkan bahwa komparasi dengan algoritma lain seperti Naïve Bayes juga dapat memberikan kontribusi yang signifikan. Hasil penelitian ini mengusulkan penggunaan algoritma C4.5 untuk mengoptimalkan distribusi dana PIP, membantu pemerintah daerah mengalokasikan sumber daya secara lebih efektif, dan menyediakan dukungan pendidikan yang tepat sasaran, mengurangi ketidaksetaraan pendidikan di Koorwilbidikcam GregeD.

2. TINJAUAN PUSTAKA

2.1. Penelitian Terdahulu

Penelitian yang dilakukan oleh Weni Ratna Sari Oktapia Ningse, S Sumarno, Zulaini Masruro Nasution Dalam jurnal “Klasifikasi Algoritma C4.5 untuk Penentuan Penerima Program Indonesia Pintar pada MIS Al-Khoirot” penelitian yang dilakukan Analisis data siswa dengan menggunakan algoritma C4. 5 untuk mengidentifikasi penerima manfaat PIP (Program Indonesia Pintar) di MIS Al-Koirot, maka dapat disimpulkan bahwa menentukan kesesuaian data siswa untuk mengidentifikasi penerima PIP dapat diselesaikan dengan menggunakan teknik data mining, yang menghasilkan 8 (delapan) aturan dan memiliki tingkat akurasi 98 persen. Dengan menggunakan teknik data mining C4. 5 diharapkan dapat

mengidentifikasi penerima manfaat PIP (Program Indonesia Pintar) di MIS AL. [1]

Dalam penelitian Nova Aprilyani, Ira Zulfa, Hendri Syahputra yang berjudul “ Penerapan Algoritma Decision Tree C4.5 Untuk Model Penentuan Penerima Beasiswa Program Indonesia Pintar (PIP) Studi Kasus Sma Negeri 3 Timang Gajah” Hasil penelitian dan diskusi tentang analisis model penentuan penerima beasiswa Program Indonesia Pintar (PIP) dengan algoritma decision tree C4.5 dapat disimpulkan bahwa menggunakan salah satu algoritma decision tree, C4.5, dapat menyelesaikan masalah menentukan apakah penerima beasiswa PIP layak atau tidak. Dengan lima aturan yang dihasilkan dari pemodelan yang menggunakan Pohon Keputusan C4.5 ini, terdapat error sebesar 20% dari pohon keputusan tersebut. Atribut yang paling signifikan diantaranya Penerima KIP, Penerima KPS, Jenis Tinggal, dan Pekerjaan Ibu. Berdasarkan evaluasi dengan hasil pengujian menggunakan alat RapidMiner, dataset sebanyak 130 data uji kemudian dibagi menjadi bagian 80:20. Di antara bagian ini terdiri dari 105 data instruksi dan 25 data pengujian. Pengukuran telah dilakukan menggunakan matriks kekacauan, yang menghasilkan tingkat akurasi 72%, Precision 85%, Recall 69 % dan Sfecificity. [2]

Dalam jurnal yang ditulis oleh ADA Abdurrahman yang berjudul “Komparasi Algoritma Naïve Bayes dan C4.5 dalam Menentukan Kelayakan Bantuan KIP (Kartu Indonesia Pintar) (Studi Kasus: DKI Jakarta)” Algoritma C4. 5 merupakan teknik data mining berupa klasifikasi data. Algoritma C4. 5 menggunakan konsep perolehan informasi atau pengurangan entropi untuk memilih partisi yang optimal. Peningkatan yang membedakan algoritme C4. 5 dari ID3 mencakup kemampuannya menangani fiturnumerik, membersihkan pohon keputusan, dan memperoleh kumpulan aturan. Proses penerapan metode pohon keputusan (C4. 5). [3]

2.2. Program Indonesia Pintar

Program Indonesia Pintar (PIP) adalah program bantuan sosial bersyarat (transfer uang syarat) di bidang pendidikan yang diberikan kepada anak-anak usia sekolah dari keluarga kurang mampu. PIP berlaku untuk anak-anak yang terdaftar di sekolah, madrasah, pesantren, kelompok belajar (mengejar paket a/b/c), dan fasilitas atau kursus pelatihan. Program ini bertujuan untuk membantu siswa memenuhi kebutuhan finansial mereka sehingga mereka dapat melanjutkan studi mereka hingga mereka lulus. Program ini merupakan modifikasi dari program sebelumnya yang dikenal sebagai Beasiswa Siswa Miskin (BSM), yang beroperasi sejak 2014. Untuk membantu siswa dari keluarga miskin atau rentan membiayai pendidikan mereka, program bantuan pemerintah yang dikenal sebagai Program Indonesia Pintar (PIP) menawarkan uang tunai, perluasan akses, dan kesempatan untuk belajar. PIP bertujuan untuk mencegah siswa putus sekolah dan mendorong mereka

untuk melanjutkan pendidikan, meringankan biaya pendidikan pribadi langsung maupun tidak langsung, dan membantu anak usia sekolah dari keluarga miskin, renta, atau prioritas mendapatkan layanan pendidikan hingga menyelesaikan pendidikan menengah, baik melalui jalur pendidikan formal maupun nonformal. dalam berbagai situasi lainnya. [4]

2.3. Algoritma C4.5

Algoritma C4.5 merupakan metode pembangunan pohon keputusan yang memberikan kemudahan dalam pemecahan masalah dan memungkinkan pengambil keputusan untuk lebih mudah memahami solusi dari masalah yang dihadapi. Tahapan dari algoritma C4.5 melibatkan perhitungan nilai Entropy untuk mengukur tingkat ketidakpastian atau kekacauan dalam dataset. Entropy dihitung sebelum dan setelah pemisahan dataset oleh suatu atribut untuk mengevaluasi seberapa baik atribut tersebut dapat mengklasifikasikan data. Berikut rumus perhitungan nilai *entropy*:

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

Keterangan:

S: himpunan kasus

n: jumlah partisi S

pi: proporsi dari Si terhadap S

Selanjutnya, perhitungan nilai Gain Ratio dilakukan untuk setiap atribut, dengan atribut yang memiliki Gain Ratio tertinggi dianggap paling informatif dan dipilih sebagai akar (root) dari pohon keputusan. Proses ini diulang untuk membangun cabang-cabang selanjutnya pada pohon keputusan hingga tidak ada lagi peningkatan informasi yang signifikan. Berikut adalah rumus perhitungan Gain Ratio:

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

Keterangan:

S : himpunan kasus

A : atribut

n : jumlah partisi atribut A

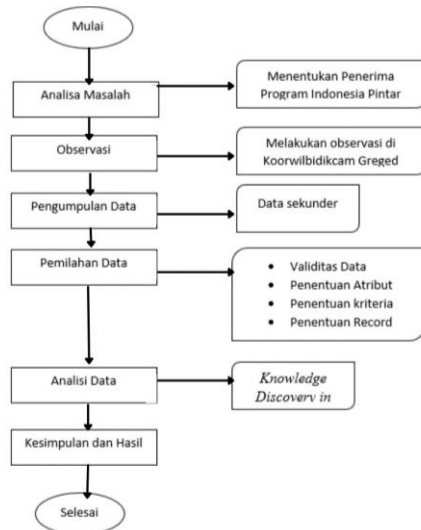
|Si| : jumlah kasus pada partisi ke-i

|S| : jumlah kasus dalam S

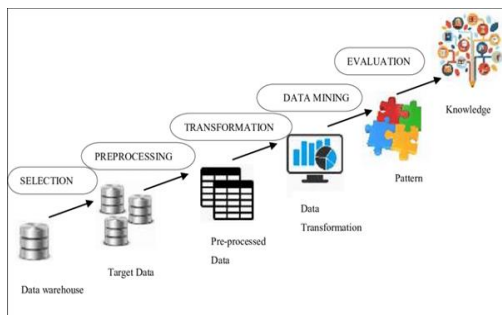
Untuk membangun Decision Tree menggunakan algoritma C4.5, tahapan dimulai dengan pemilihan atribut sebagai akar berdasarkan perhitungan jumlah kasus dan target atribut. Entropy dihitung untuk menilai informativitas suatu atribut dalam menghasilkan atribut output. Informasi gain digunakan untuk memisahkan objek, dan atribut dengan Gain Ratio tertinggi dipilih sebagai cabang berikutnya. Proses ini diulang hingga tidak ada lagi peningkatan informasi yang signifikan atau semua atribut telah digunakan. Algoritma C4.5 memberikan

landasan matematis yang kuat melalui perhitungan entropy dan gain untuk membangun pohon keputusan yang informatif dan efektif dalam mengklasifikasikan data.

3. METODE PENELITIAN



Metode dalam penelitian ini menggunakan Knowledge Discovery in Databases (KDD) yang terdiri dari Data Selection, Data Preprocessing, Data transformation, Data mining, Pattern evolution, knowledge, seperti terlihat gambar 1.



Gambar 2. Tahapan KDD

Tahapan-tahapan KDD mencakup serangkaian langkah-langkah yang harus diambil untuk mengubah data mentah menjadi pengetahuan yang dapat diinterpretasi. Berikut adalah tahapan-tahapan dalam proses KDD:

3.1. Data

Data jumlah siswa yang berada di Sekolah Dasar (SD) di Koorwilbidikcam Greged didapatkan dari data yang bersumber dari Aplikasi Dapodik Versi 2024.

Penelitian ini menggunakan dataset sebanyak 3000 dataset.

3.2. Data Selection

Pada tahap data selection, dilakukan pemilihan data yang akan digunakan dalam proses prediksi menggunakan tools RapidMiner. Seleksi data ini memiliki tujuan untuk memilih data yang relevan dalam proses klasifikasi, dan tahapan seleksi dilakukan di Excel. Dalam penelitian ini, data yang digunakan berasal dari jumlah keadaan data kelas 1 hingga kelas 6 di Koorwilbidikcam Greged untuk melakukan klasifikasi penerima Program Indonesia Pintar. Hasil seleksi data menunjukkan adanya 5 atribut yang relevan, yaitu pekerjaan orangtua, penghasilan orangtua, penerima KPS, pemegang SKTM, dan alat transportasi. Sementara itu, atribut seperti Nama, NIPD, Jenis Kelamin, NISN, dan Alamat tidak digunakan dalam proses ini. Detail atribut yang terpilih dan tidak terpakai dapat ditemukan dalam Tabel 1 dan Tabel 2. Seleksi ini memastikan hanya atribut yang relevan dan berkontribusi pada proses klasifikasi yang dipertahankan, memfokuskan analisis pada informasi yang lebih signifikan.

Tabel 1. Tabel Sebelum pemilihan atribut

No	Atribut
1	Jenis Tinggal
2	Alat Transportasi
3	Pekerjaan orangtua
4	Penghasilan orangtua
5	Penerima KPS
6	Pemegang SKTM
7	Rombel Saat Ini
8	Alat Transportasi

Tabel 2. Tabel Sesudah pemilihan atribut

No	Atribut
1	Pekerjaan orangtua
2	Penghasilan orangtua
3	Penerima KPS
4	Pemegang SKTM
5	Alat Transportasi

3.3. Data Preprocessing

Pada tahap preprocessing data, kesalahan adalah untuk mengidentifikasi dan memperbaiki berbagai elemen data, termasuk kesalahan (seperti pengurangan ukuran teks), keakuratan, relevansi atau informasi yang tidak relevan, dan informasi yang hilang. Hasil pemilihan data seperti ditampilkan pada tabel 3

Tabel 3. Data Preprocessing

No	Pekerjaan orangtua	Penghasilan orangtua	Alat Transportasi	Penerima KPS	Pemegang SKTM
1	Pedagang	Kurang dari Rp. 500,000	Jalan kaki	Tidak	Tidak
2	Buruh	Kurang dari Rp. 500,000	Jalan kaki	Tidak	Ya

No	Pekerjaan orangtua	Penghasilan orangtua	Alat Transportasi	Penerima KPS	Pemegang SKTM
3	Buruh	Rp. 1,000,000 - Rp. 1,999,999	Jalan kaki	Tidak	Ya
4	Buruh	Kurang dari Rp. 500,000	Jalan kaki	Tidak	Ya
5	Buruh	Kurang dari Rp. 500,000	Jalan kaki	Tidak	Ya
6	Buruh	Kurang dari Rp. 500,000	Jalan kaki	Tidak	Ya
7	Buruh	Rp. 500,000 - Rp. 999,999	Jalan kaki	Tidak	Tidak
8	Buruh	Rp. 1,000,000 - Rp. 1,999,999	Jalan kaki	Ya	Ya
9	PNS/TNI/Polri	Rp. 2,000,000 - Rp. 4,999,999	Jalan kaki	Tidak	Tidak
10	Buruh	Rp. 1,000,000 - Rp. 1,999,999	Jalan kaki	Tidak	Ya
2999	Pedagang	Kurang dari Rp. 500,000	Jalan kaki	Tidak	Ya
3000	Pedagang	Kurang dari Rp. 500,000	Jalan kaki	Tidak	Ya

3.4. Data Transformation

Tahapan data transformation dalam hasil pembahasan melibatkan perubahan atribut data dari ekstensi CSV ke dalam format yang sesuai untuk diproses dalam perangkat lunak RapidMiner. Transformasi ini diilustrasikan dalam Gambar 2, di mana operator set role digunakan untuk mengubah atribut "name" menjadi "Pemegang SKTM" dan menetapkan atribut "label" sebagai target role. Data transformation memiliki fungsi utama dalam proses ini, yaitu mengubah, membersihkan, dan menyusun data ke dalam format yang dapat digunakan untuk analisis dan mendukung pengambilan keputusan.

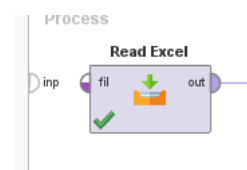
Row No.	Pemegang	Alat Transp.	Penerima K.	Penghasilan O.	Penghasilan
1	Tidak	Jalan kaki	Tidak	Lainnya	Kurang dari ...
2	Ya	Jalan kaki	Tidak	Lainnya	Kurang dari ...
3	Ya	Jalan kaki	Tidak	Pedagang Ka...	Rp. 1,000,00...
4	Ya	Jalan kaki	Tidak	Lainnya	Kurang dari ...
5	Ya	Jalan kaki	Tidak	Lainnya	Kurang dari ...
6	Ya	Jalan kaki	Tidak	Lainnya	Kurang dari ...
7	Tidak	Jalan kaki	Tidak	Buruh	Rp. 500,000...
8	Ya	Jalan kaki	Ya	Tidak bekerja	Tidak Berpen...
9	Tidak	Jalan kaki	Tidak	Wiraswasta	Rp. 2,000,00...
10	Ya	Jalan kaki	Tidak	Tidak bekerja	Tidak Berpen...
11	Tidak	Jalan kaki	Tidak	Lainnya	Rp. 500,000...
12	Tidak	Jalan kaki	Tidak	Lainnya	Rp. 500,000...
13	Ya	Jalan kaki	Tidak	Buruh	Rp. 1,000,00...
14	Tidak	Jalan kaki	Tidak	Buruh	Rp. 1,000,00...
15	Ya	Jalan kaki	Tidak	Lainnya	Kurang dari ...
16	Ya	Jalan kaki	Tidak	Lainnya	Kurang dari ...
17	Ya	Jalan kaki	Tidak	Lainnya	Kurang dari ...
18	Ya	Jalan kaki	Tidak	Lainnya	Kurang dari ...
19	Tidak	Dipenda motor	Tidak	Wiraswasta	Rp. 2,000,00...

Gambar 3. Data Transformation

Hasil dari proses transformasi ini tercermin pada Gambar 2, di mana atribut yang ditetapkan dan disusun sesuai kebutuhan analisis, memastikan bahwa data siap untuk digunakan dalam langkah-langkah selanjutnya dalam proses klasifikasi menggunakan RapidMiner.

3.5. Data Mining

Tahap selanjutnya setelah proses transformasi data adalah implementasi model algoritma C4.5 (Decision Tree) melalui RapidMiner yang berfungsi sebagai untuk membangun pohon keputusan yang dapat digunakan untuk klasifikasi atau prediksi berdasarkan data yang diberikan., adapun tahapan dalam proses Data Mining adalah sebagai berikut:



Gambar 4. Tampilan import data excel

Penjelasan Gambar 4. adalah tampilan panel proses, selanjutnya yaitu buka Panel operators cari Read Excel, karena tipe file yang digunakan dalam penelitian ini menggunakan format ekstensi Excel, kemudian drag dan drop read excel tersebut kedalam panel proses.

4. HASIL DAN PEMBAHASAN

4.1. Hasil

a. Importing Data

Data yang digunakan dalam penelitian ini hanya penghasilan orang tua, pekerjaan orang tua, alat transportasi dari rumah ke sekolah dan status penerima KPS. Dataset ditampilkan pada gambar 5.

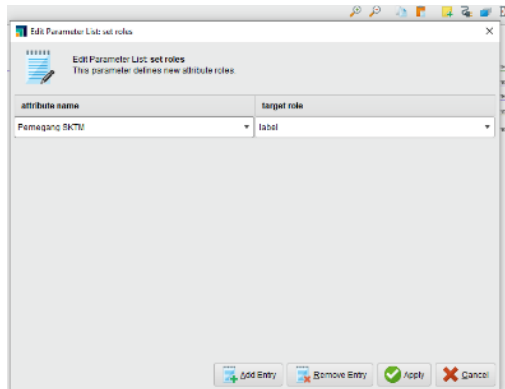
Row No.	Pemegang	prediction	confidence	confidence	Penerima K.	Penghasilan	Alat Transp.	Penerima K.
1	Ya	Ya	0.875	0.125	Buruh	Kurang dari ...	Jalan kaki	Tidak
2	Ya	Ya	0.875	0.125	Buruh	Kurang dari ...	Jalan kaki	Tidak
3	Ya	Ya	0.954	0.046	Wiraswasta	Rp. 1,000,00...	Jalan kaki	Tidak
4	Ya	Ya	0.981	0.019	Buruh	Rp. 1,000,00...	Jalan kaki	Tidak
5	Ya	Ya	0.954	0.046	Wiraswasta	Rp. 1,000,00...	Jalan kaki	Tidak
6	Ya	Ya	0.981	0.019	Buruh	Rp. 1,000,00...	Jalan kaki	Tidak
7	Tidak	Ya	0.735	0.265	Pedagang	Rp. 1,000,00...	Dipenda motor	Tidak
8	Ya	Ya	0.735	0.265	Pedagang	Rp. 1,000,00...	Jalan kaki	Tidak
9	Ya	Ya	0.981	0.019	Buruh	Rp. 1,000,00...	Jalan kaki	Tidak
10	Ya	Ya	0.735	0.265	Pedagang	Rp. 1,000,00...	Jalan kaki	Tidak
11	Ya	Ya	0.735	0.265	Pedagang	Rp. 1,000,00...	Jalan kaki	Tidak
12	Ya	Ya	0.930	0.070	Pedagang	Kurang dari ...	Jalan kaki	Tidak
13	Ya	Ya	0	0	Buruh	Rp. 2,000,00...	Jalan kaki	Ya
14	Ya	Ya	0.930	0.070	Pedagang	Kurang dari ...	Jalan kaki	Tidak
15	Tidak	Ya	0.930	0.070	Pedagang	Kurang dari ...	Jalan kaki	Ya
16	Ya	Ya	0.930	0.070	Pedagang	Kurang dari ...	Jalan kaki	Tidak
17	Ya	Ya	0.981	0.019	Buruh	Rp. 1,000,00...	Jalan kaki	Tidak
18	Ya	Ya	0.981	0.019	Buruh	Rp. 1,000,00...	Jalan kaki	Tidak
19	Ya	Ya	0.981	0.019	Buruh	Rp. 1,000,00...	Jalan kaki	Tidak

Gambar 5 Dataset

b. Mengubah atribut name dan target role dengan menggunakan operator Set Role

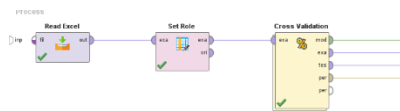


Gambar 6. Operator set rol



Gambar 7. Parameter operator set role

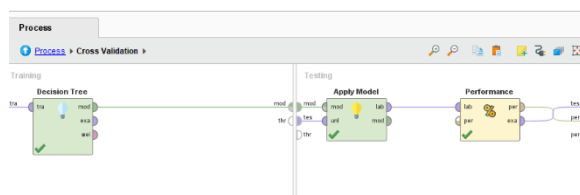
- c. Menghubungkan port Set Role dengan Cross Validation



Gambar 8. port set role dan port Cross Validation

Penjelasan Gambar 7 adalah penggunaan operator set role dengan Cross Validation menentukan atribut mana yang akan digunakan sebagai label (atribut yang nilainya ingin diprediksi) dalam model algoritma C4.5.

- d. Melakukan Data Training dan Data Testing

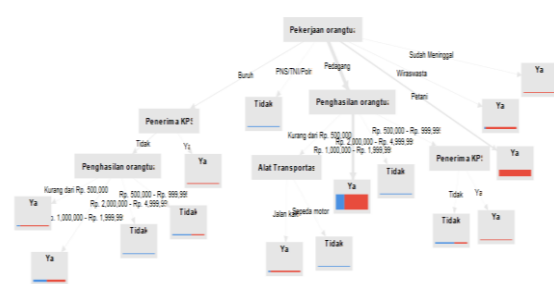


Gambar 9. Proses Model

Penjelasan gambar 8 yaitu Data dibagi menjadi dua bagian, didalam data training terdapat operator algoritma Decision Tree yaitu model menghasilkan model pohon keputusan yang dapat digunakan untuk klasifikasi.

- e. Tampilan Decision Tree

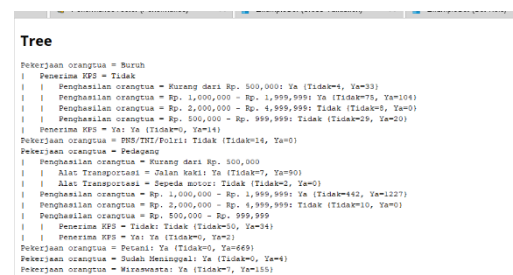
Setelah menghitung dan menguji data untuk setiap atribut menggunakan algoritma C4.5 maka akan di dapatkan pola pohon keputusan akhir yang ditampilkan pada gambar 9.



Gambar 10. Hasil Decision Tree

Penjelasan gambar 9 yaitu hasil pohon keputusan dimana atribut pekerjaan orangtua merupakan root node yang mewakili seluruh dataset. Kemudian terdapat percabangan atribut penghasilan orangtua, alat transportasi, dan penerima KPS yang kemudian hasil pohon tersebut menunjukan label Ya dan Tidak dalam klasifikasi menggunakan algoritma C4.5.

- f. Tampilan Deskripsi Decision Tree



Gambar 11. Tampilan Deskripsi Decision Tree

Penjelasan gambar 10 yaitu Hasil yang didapat dari algoritma C4 adalah gambaran pohon keputusan secara lengkap. Hasil yang diuraikan juga menunjukkan bahwa penggunaan algoritma data mining C4. 5 sangat direkomendasikan untuk digunakan dalam proses penambangan data yang menghasilkan banyak hasil dan menampilkan dalam pohon keputusan.

- g. Tingkat Akurasi

	true Tidak	true Ya	class precision
pred Tidak	112	56	66.67%
pred Ya	536	2296	81.07%
class recall	17.28%	97.62%	

Gambar 12. Tingkat Akurasi

Penjelasan gambar 11 yaitu Berdasarkan pengolahan data menggunakan software rapidminer didapat nilai akurasi sistem sebesar 80.27 % artinya bahwa rule yang dihasilkan tingkat kebenarannya mendekati 100%. Dimana model yang telah dibentuk diuji tingkat akurasinya dengan memasukkan atau uji yang berasal dari data training dengan menggunakan

cross validation pada aplikasi rapidminer untuk menguji tingkat akurasi.

h. Recall

	true Tidak	true Ya	class precision
pred. Tidak	112	56	66.67%
pred. Ya	536	2296	81.07%
class recall	17.28%	97.62%	

Gambar 13. Recall Data

Penjelasan gambar 12 yaitu Dari hasil pengukuran data diatas nilai recall yang diperoleh yaitu sebesar 57.45%. Dari tabel tersebut diketahui prediksi Tidak dengan true Tidak mencapai 112 data dan true Ya sebanyak 56 data, dengan hasil pencapaian nilai presisi sebesar 66.67%. Sedangkan untuk prediksi Ya untuk true Tidak mencapai 536 data dan untuk true Ya sebanyak 2296 data dengan hasil pencapaian nilai presisi sebesar 81.07%. Untuk recall data responden yang sangat tinggi terdapat pada class recall true Ya mencapai 97.62% sedangkan untuk class recall true tidak mencapai 17.28%.

i. Class Precision

	true Tidak	true Ya	class precision
pred. Tidak	112	56	66.67%
pred. Ya	536	2296	81.07%
class recall	17.28%	97.62%	

Gambar 14. Class Precision Data

Penjelasan gambar 13 yaitu Dari hasil pengukuran data diatas nilai precision yang diperoleh yaitu sebesar 73.87%. Dari tabel tersebut diketahui prediksi Tidak dengan true Tidak mencapai 112 data dan true Ya sebanyak 56 data, dengan hasil pencapaian nilai presisi sebesar 66.67%. Sedangkan untuk prediksi Ya untuk true Tidak mencapai 536 data dan untuk true Ya sebanyak 2296 data dengan hasil pencapaian nilai presisi sebesar 81.07%. Untuk recall data responden yang sangat tinggi terdapat pada class recall true Ya mencapai 97.62% sedangkan untuk class recall true tidak mencapai 17.28%.

Algoritma C4.5 berhasil diimplementasikan untuk menghasilkan pohon keputusan, seperti yang terlihat pada Gambar 4.6. Pohon keputusan ini memberikan wawasan tentang faktor-faktor dominan dalam menentukan penerima Program Indonesia Pintar (PIP). Faktor-faktor yang paling berpengaruh, seperti terlihat pada hasil Decision Tree, melibatkan pekerjaan orang tua, penghasilan orang tua, dan status penerima Kartu Perlindungan Sosial (KPS). Hasil pengujian dari data training dan data testing

menunjukkan tingkat akurasi sebesar 80.27%, recall sebesar 57.45%, dan precision sebesar 73.96%. Decision Tree juga menggambarkan bahwa faktor pekerjaan orang tua dan penghasilan orang tua memiliki pengaruh terbesar terhadap keputusan penerimaan PIP. Sebagai contoh, jika pekerjaan orang tua adalah buruh dan menerima KPS, atau pekerjaan sebagai buruh tanpa KPS namun dengan penghasilan kurang dari Rp. 500.000,00, maka siswa layak menerima PIP. Namun, jika pekerjaan orang tua adalah PNS/TNI/Polri, atau pekerjaan sebagai pedagang dengan penghasilan di atas Rp. 4.999.999, maka siswa tidak layak menerima PIP. Berbagai kondisi seperti ini dijelaskan dalam hasil Decision Tree untuk membantu pengambilan keputusan terkait penerimaan PIP.

4.2. Pembahasan

Dalam penelitian ini dilakukan pengumpulan data yang bersumber dari Aplikasi Dapodik Versi 2024. Kemudian dilakukan pembersihan data yang kemudian data tersebut ditransformasi dengan atribut-atribut yang dapat memengaruhi kinerja algoritma C4.5. Penelitian ini menggunakan algoritma C4.5 untuk menghasilkan Decision Tree sebagai output. Data yang diperoleh diimpor ke RapidMiner melalui operator Read Excel, lalu dibagi menjadi dua bagian yaitu data training dan data testing. Evaluasi model dilakukan dengan menggunakan Cross Validation. Dimana dataset dibagi menjadi 80 % data Training dan 20% data Testing. Tahap selanjutnya adalah membangun model menggunakan data training. Model klasifikasi Decision Tree digunakan untuk memprediksi penerima PIP dengan mengklasifikasikan data ke dalam kelas "Ya" dan "Tidak".

5. KESIMPULAN DAN SARAN

Penelitian ini mengeksplorasi penerapan Algoritma C4.5 dalam klasifikasi penentuan penerima Program Indonesia Pintar (PIP) di Koorwilbidikcam Greged. Hasil penelitian menunjukkan bahwa Algoritma C4.5 efektif dalam melakukan klasifikasi pada dataset penentuan penerima PIP dengan tingkat akurasi dan kinerja model yang memuaskan. Proses seleksi atribut pada algoritma ini membantu mengidentifikasi atribut kritis yang memiliki kontribusi signifikan dalam penentuan penerima PIP, mempermudah pemahaman faktor-faktor yang memengaruhi keputusan. Model klasifikasi yang dihasilkan memberikan pemahaman lebih lanjut terhadap faktor-faktor yang mempengaruhi keputusan di wilayah Koorwilbidikcam Greged. Untuk pengembangan lebih lanjut, disarankan untuk menggali potensi model klasifikasi yang lebih kompleks, menambah variabel atau data tambahan, dan melakukan penelitian terkait optimalisasi hyperparameter pada Algoritma C4.5 untuk mencapai model yang lebih optimal dan efisien.

DAFTAR PUSTAKA

- [1] Irsyad Zamjani, Herlinawati, N. S. Perdana, F. Widiaputera, and S. N. Azizah, *Biaya Satuan & Lini Masa Pengelolaan Program Indonesia Pintar*. 2020.
- [2] W. Ningse, S. Sumarno, and Z. M. Nasution, "Klasifikasi Algoritma C4. 5 untuk Penentuan Penerima Program Indonesia Pintar pada MIS Al-Khoirot C4. 5 Algorithm Classification for Determining Smart ...," download.garuda.kemdikbud.go.id. [Online]. Available: <http://download.garuda.kemdikbud.go.id/article.php?article=2998597&val=27036&title=C45> Algorithm Classification for Determining Smart Indonesia Program Recipients at MIS Al-Khoirot
- [3] H. S. Nova Aprilyani, Ira Zulfa, Penerapan Algoritma Decision Tree C4.5 Untuk Model Penentuan Penerima Beasiswa Program Indonesia Pintar (Pip) Studi Kasus Sma Negeri 3 Timang Gajah. Takengon: Jurnal Teknik Informatika Dan Elektro, 2023.
- [4] ADA Abdurrahman, "Komparasi Algoritma Naïve Bayes dan C4.5 dalam Menentukan Kelayakan Bantuan KIP (Kartu Indonesia Pintar) (Studi Kasus: DKI Jakarta)," 2020, [Online]. Available: <https://repository.mercubuana.ac.id/id/eprint/68157>
- [5] C. R. and T. Secretary General of the Ministry of Education, "Peraturan Sekretaris Jenderal Kementerian Pendidikan, Kebudayaan, Riset, and Teknologi Nomor 14 Tahun 2022 tentang Petunjuk Pelaksanaan Program Indonesia Pintar Pendidikan Dasar dan Pendidikan Menengah (Regulation of the Secretary General of the Ministry," Jar. Dokumentasi dan Inf. Huk. Nas. [JDIHN], 2022, [Online]. Available: <https://jdih.kemdikbud.go.id/>
- [6] D. S. Eko Budiarto, Rino, Susanto Hariyanto, Penerapan Data Mining Untuk Rekomendasi Beasiswa Pada SD Maria Mediatrix Menggunakan Algoritma C4.5. 2022. [Online]. Available: <https://doi.org/10.31253/algov3i2.1019%0A>
- [7] P. P. Haryoto, H. Okprana, and ..., "Algoritma C4. 5 Dalam Data Mining Untuk Menentukan Klasifikasi Penerimaan Calon Mahasiswa Baru," TIN Terap. Inform. ..., 2021, [Online]. Available: <http://ejurnal.seminar-id.com/index.php/tin/article/view/919>
- [8] N. W. O. Pratiwi, N. W. Utami, and I. Putra, "Klasifikasi Penentuan Penerima Bantuan Sosial Tunai (BST) Menggunakan Algoritma C4. 5 Di Desa Keramas, Gianyar, Bali," J. Inform. Teknol. ..., 2022, [Online]. Available: <http://www.jurnal.uts.ac.id/index.php/JINTEKS/article/view/1667>
- [9] N. Hidayah, SISTEM KLASIFIKASI PENERIMA BERAS MISKIN MENGGUNAKAN ALGORITMA DECISION TREE C4. 5 (Studi Kasus: Kelurahan Tambakmerang, Girimarto eprints.uty.ac.id, 2019. [Online]. Available: <http://eprints.uty.ac.id/4136/>
- [10] N. Nurahman, "KLASIFIKASI PENERIMA BANTUAN SOSIAL DI DESA BATUAH MENGGUNAKAN METODE ALGORITMA C4. 5," J. Tekinkom (Teknik Inf. dan ..., 2022, [Online]. Available: <http://jurnal.murnisadar.ac.id/index.php/Tekinkom/article/view/516>