

IMPLEMENTASI ALGORITMA C4.5 UNTUK PREDIKSI INDEKS KEDALAMAN KEMISKINAN DAN INDEKS KEPARAHAN KEMISKINAN MENURUT DAERAH

Iqbal Baihaqi¹, Anis Cahayasih², Al Fathya Khaerunisa³,
Mourizka Chaerunnisa⁴, Muhamad Arya Wildana⁵, Baginda Oloan Lubis⁶

^{1,2,3,4,5} Sistem Informasi, Bina Sarana Informatika

⁶ Rekayasa Perangkat Lunak, Bina Sarana Informatika

Jl. Kramat Raya No. 98 Jakarta, Indonesia

⁴ mourizkac@gmail.com

ABSTRAK

Kemiskinan merupakan salah satu tantangan terbesar yang dihadapi Indonesia dalam upaya mencapai pembangunan berkelanjutan. Penelitian ini bertujuan untuk menganalisis tingkat dan dinamika kemiskinan di Indonesia dengan pendekatan *Multidimensional Poverty Index* (MPI). Data yang digunakan meliputi indikator-indikator kemiskinan seperti pendidikan, kesehatan, dan standar hidup yang diperoleh dari survei sosial ekonomi nasional (Susenas). Hasil penelitian menunjukkan bahwa meskipun terjadi penurunan tingkat kemiskinan moneter dalam dua dekade terakhir, kemiskinan multidimensi masih cukup tinggi di beberapa wilayah, terutama di kawasan timur Indonesia. Faktor-faktor seperti akses terhadap layanan kesehatan, pendidikan yang rendah, dan ketimpangan antarwilayah menjadi kontributor utama. Studi ini juga memberikan rekomendasi kebijakan, termasuk peningkatan akses infrastruktur dasar dan penguatan program perlindungan sosial untuk mempercepat penurunan kemiskinan. Kebijakan meliputi peningkatan akses infrastruktur, pemerataan sumber daya, dan penguatan program perlindungan sosial. Studi ini diharapkan dapat memberikan kontribusi dalam merumuskan strategi Pembangunan. Temuan ini diharapkan dapat menjadi dasar dalam perumusan strategi pembangunan yang lebih inklusif dan berkelanjutan di masa depan.

Kata kunci : Kemiskinan di Indonesia, *Multidimensional Poverty Index* (MPI), ketimpangan, pembangunan berkelanjutan.

1. PENDAHULUAN

Kemiskinan adalah situasi dimana seseorang atau rumahtangga mengalami kesulitan untuk memenuhi kebutuhan dasar, sementara lingkungan penduduknya kurang memberikan peluang untuk meningkatkan kesejahteraan secara berkesinambungan atau untuk keluar dari kerentanan [1].

Masalah kemiskinan di Indonesia telah menjadi fokus utama dalam proses pembangunan nasional. Kemiskinan tidak hanya terbatas pada daerah pedesaan, tetapi juga terjadi di perkotaan. Masalah ini telah ada sejak era reformasi Indonesia dan merupakan tantangan kompleks yang dihadapi oleh semua pemerintahan di dunia, termasuk Indonesia [2].

Indonesia merupakan salah satu Negara berkembang di Asia khususnya Asia Tenggara. Salah satu masalah yang sering dihadapi oleh negara berkembang adalah Kemiskinan. Bagi pemerintah Indonesia masalah kemiskinan merupakan masalah lama yang sulit untuk diselesaikan. Secara teoritis, telah dilakukan upaya dalam pengentasan kemiskinan pada suatu daerah dengan upaya pertumbuhan ekonomi yang berkualitas dengan beberapa program kerja pemerintah untuk meningkatkan pertumbuhan ekonomi yang berkualitas [3].

Persoalan kemiskinan merupakan salah satu permasalahan pokok yang dihadapi bangsa Indonesia sejak dulu hingga sekarang. Berbagai perencanaan,

kebijakan serta program pembangunan yang telah dan akan dilaksanakan pada intinya adalah mengurangi jumlah penduduk miskin [4].

Faktor lain yang menyebabkan ketimpangan pendapat dalam suatu daerah salah satunya adalah kemiskinan. Kemiskinan sering dipahami sebagai keadaan kekurangan uang dan barang untuk menjamin kelangsungan hidup [5].

Pengetasan kemiskinan di wilayah tersebut diharapkan mampu mengatasi kesenjangan pertumbuhan perekonomian masyarakat miskin. Terkait pengetasan kemiskinan, Pemerintah mempunyai berbagai macam program untuk mengatasi kemiskinan. Akan tetapi, terdapat berbagai kendala dalam penyaluran bantuan dari pemerintah untuk mengklasifikasikan para calon penerima bantuan. Sehingga, membuat proses penyaluran bantuan berjalan lama dan tidak tepat sasaran. Pengetasan kemiskinanpun juga belum bisa dikendalikan. Pada pengumpulan data penduduk yang sudah ada di Desa, dengan memanfaatkan teknik data mining pengklasifikasian dapat ditentukan [6].

Penelitian dari (Prakoso et al. 2019) menyebutkan bahwa kemandirian keuangan daerah berpengaruh negatif signifikan terhadap kemiskinan [7].

BPS menegaskan bahwa kemiskinan menandakan ketidakmampuan individu dalam memenuhi kebutuhan dasar yang esensial untuk hidup yang layak, baik dalam bentuk bahan pangan

maupun non-pangan [2].

Berlandaskan infomasi yang diterima dari Badan Pusat Statistik (BPS) tahun 2023, presentase penduduk miskin pada tahun 2023 sebesar 14,90 persen, menurun 0,30 persen poin dibanding tahun 2022. Jumlah penduduk miskin pada tahun 2023 sebanyak 138,99 ribu orang, menurun 2,26 ribu dibandingkan tahun 2022 [5].

Menurut BPS (2016), kemiskinan adalah keadaan tidak mampu, secara finansial atau lainnya untuk memenuhi kebutuhan dasar fisik dan nutrisi seseorang [8].

Adapun yang menjadi konsep kemiskinan ada tiga yaitu :

- a. Kemiskinan absolut dirumuskan dengan membuat ukuran tertentu konkret, ukuran ini lazimnya berorientasi pada kebutuhan hidup dasar minimum anggota masyarakat yang dipergunakan sebagai acuan memang berlainan
- b. Kemiskinan relatif dirumuskan dengan demensi tepat dan waktu. Asumsinya adalah kemiskinan disuatu daerah berbeda dengan daerah lainnya, konsep kemiskinan ini lazimnya diukur berdasarkan pertimbangan anggota masyarakat tertentu, dengan berorientasi pada derajat kekayaan hidup. Konsep ini juga telah memperoleh banyak keritikan terutama sangat sulit menentukan bagaimana hidup yang layak itu. Ukuran kelayakan juga beragam dan terus berubah-ubah
- c. Kemiskinan subyektif dirumuskan berdasarkan kelompok kemiskinan itu sendiri. Konsep ini tidak mengenal dan tidak memperhitungkan. Kelompok menurut ukuran kita berbeda dibawah kemiskinan, oleh karena itu konsep kemiskinan ini dianggap lebih tepat apabila dipergunakan untuk memahami kemiskinan dan merumuskan cara atau strategi yg efektif untuk penanggulangannya [1].

Badan Pusat Statistik (BPS) sebagai lembaga statistik resmi di Indonesia memiliki peran vital dalam menyediakan data kemiskinan yang akurat dan terpercaya. Data BPS menjadi acuan utama bagi pemerintah dalam merumuskan kebijakan dan program penanggulangan kemiskinan, serta bagi para peneliti dan akademisi dalam melakukan penelitian terkait kemiskinan. Meskipun data kemiskinan BPS banyak digunakan oleh berbagai pihak, terdapat kesenjangan pemahaman antara BPS dan pengguna data terkait metodologi pengolahan data. Kesenjangan ini dapat menimbulkan misinterpretasi, misinformasi, ketidakpercayaan terhadap data kemiskinan, dan pada akhirnya menghambat upaya penanggulangan kemiskinan [9].

Menurut Chamsyah penyebab kemiskinan adalah berupa hubungan-hubungan kompleks antara individu yang hidup dengan daya lemah dalam suatu tradisi keluarga, masyarakat dalam ruang struktur sosial (negara) yang rumit dan menganut

sistem modern dalam cara produksinya. Individu sebagai penyebab kemiskinan menempati posisi pertama dalam kondisi kurang sejahtera. Menurut Munker dan Walter, kemiskinan ditimbulkan oleh faktor-faktor penyebab dan dimensi-dimensi yang berbeda-beda. Hal ini tergantung pada kelompok yang miskin tinggal di pedesaan atau perkotaan, dan apakah mereka hidup dalam struktur keluarga kecil, keluarga inti atau seorang diri [1].

Penyebab kemiskinan menurut Suharto (2005) yaitu: (a) faktor individual, terkait dengan aspek patologis, termasuk kondisi fisik dan psikologis individu yang miskin, (b) faktor sosial, kondisi-kondisi lingkungan sosial yang menjebak seseorang menjadi miskin. Misalnya diskriminasi berdasarkan usia, gender, etnis yang menyebabkan seseorang menjadi miskin, (c) faktor kultural, kondisi atau kualitas budaya yang menyebabkan kemiskinan, (d) faktor struktural, berkaitan dengan struktur atau sistem yang tidak adil, tidak sensitif dan tidak accessible sehingga menyebabkan seseorang menjadi miskin [1].

Dua strategi yang harus diambil dalam upaya mengatasi kemiskinan adalah, pertama, memberikan perlindungan kepada keluarga dan kelompok masyarakat miskin dengan memenuhi kebutuhan mereka di berbagai sektor. Kedua, memberikan pelatihan kepada mereka agar memiliki keterampilan yang diperlukan untuk mencegah timbulnya kemiskinan baru. Langkah-langkah ini diimplementasikan dengan tujuan untuk mencapai cita-cita bangsa, yaitu menciptakan masyarakat yang adil dan sejahtera [10].

2. TINJAUAN PUSTAKA

2.1. Algoritma C4.5 dalam Data Mining

Data mining dapat diterapkan pada segala bentuk data atau pemrosesan informasi berskala besar, serta pada berbagai aplikasi *computer decision support system* (DSS), termasuk kecerdasan buatan (AI)[11].

Untuk mengimplementasikan algoritma C4.5 dalam prediksi Indeks Kedalaman Kemiskinan dan Indeks Keparahan Kemiskinan menurut daerah, kita akan membahas langkah-langkah yang lebih rinci mengenai bagaimana menghitung Gain (Information Gain) dalam konteks ini, berdasarkan data yang terdapat dalam jurnal atau publikasi penelitian [12].

2.2. Langkah 1 : Menyusun Tujuan

Sebelum kita mengimplementasikan algoritma C4.5, kita perlu memahami konteks dan tujuan. Dalam hal ini, kita bertujuan untuk memprediksi Indeks Kedalaman Kemiskinan (IPK) dan Indeks Keparahan Kemiskinan (IKK) berdasarkan data yang diperoleh dari daerah tertentu. Data yang digunakan mungkin meliputi beberapa atribut seperti:

- a. Pendapatan per kapita
- b. Tingkat pendidikan
- c. Tingkat pengangguran
- d. Kondisi rumah tangga
- e. Akses ke fasilitas kesehatan

Dalam hal ini, atribut-atribut ini digunakan untuk memprediksi dua jenis indeks kemiskinan berdasarkan daerah.

2.3. Langkah 2 : Memahami Algoritma C4.5

C4.5 adalah algoritma pohon keputusan yang digunakan untuk membangun model prediktif. [13]

Proses pemilihan atribut dilakukan dengan menghitung gain (atau lebih tepatnya *Information Gain*) yang mengukur seberapa baik atribut membagi data menjadi subset yang lebih murni.

C4.5 menggunakan Entropy untuk menghitung ketidakpastian dalam data, dan Gain digunakan untuk memilih atribut terbaik [14].

2.4. Langkah 3 : Menghitung Entropy dan Gain

Untuk menghitung Gain di C4.5, kita memerlukan dua konsep utama:

- a. Entropy : Ukuran ketidakpastian atau ketidakteraturan dalam dataset.
- b. Gain : Mengukur seberapa banyak ketidakpastian dikurangi setelah data dibagi berdasarkan suatu atribut.

2.5. Pemanfaatan Entropy dan Information Gain

Entropy dan Information Gain adalah konsep kunci dalam algoritma C4.5 untuk menentukan atribut terbaik dalam membagi dataset.[15]

Entropy mengukur tingkat ketidakpastian dalam dataset, sementara Information Gain berfungsi untuk mengurangi ketidakpastian tersebut dengan memilih atribut yang memberikan pemisahan terbaik. [16]

Gabungan kedua konsep ini memastikan bahwa pohon keputusan yang dihasilkan menjadi lebih efisien dan akurat dalam proses klasifikasi.[17]

2.6. Menghitung Entropy untuk Atribut Utama

Entropy adalah ukuran ketidakpastian atau kekacauan dalam data. Dalam konteks pengambilan keputusan, terutama dalam pembelajaran mesin, seperti pohon keputusan, entropy digunakan untuk mengevaluasi tingkat impurity (ketidakmurnian) suatu dataset berdasarkan distribusi label target.[15]

Secara matematis, entropy $E(S)$ untuk dataset S dapat dihitung menggunakan rumus berikut:

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

Penjelasan:

- S : himpunan kasus
- N : jumlah partisi S
- Pi : proporsi dari Si terhadap S
- (-) : Digunakan karena logaritma probabilitas selalu negatif

2.7. Menghitung Information Gain (G)

Gain (atau Information Gain) adalah ukuran yang digunakan untuk menentukan efektivitas suatu atribut dalam memisahkan dataset menjadi subset yang lebih homogen (lebih murni). Dengan kata lain, Gain mengukur pengurangan Entropy setelah data dipisahkan berdasarkan suatu atribut tertentu.

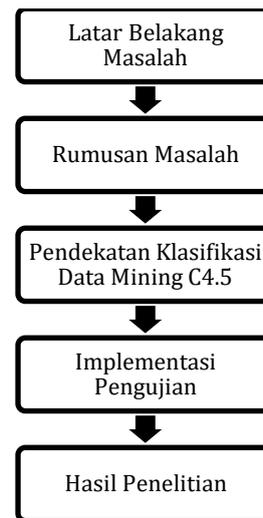
$$GAIN(S, A) = E(S) - \sum_{v \in values(A)} \frac{|S_v|}{|S|} \cdot E(S_v) \quad (2)$$

Penjelasan:

- S : himpunan data awal
- A : atribut yang di gunakan untuk membagi data
- Values : nilai-nilai unik dari atribut A
- |Sv| : subset data S yang memiliki nilai v pada atribut A
- $\frac{|S_v|}{|S|}$: proporsi data dalam subset Sv terhadap total S
- E(Sv) : entropy dari subset Sv

3. METODE PENELITIAN

Berikut merupakan tahapan penelitian pada penelitian ini:



Gambar 1. Tahapan Penelitian

Pada tahapan penelitian ini dilakukan peneliti dalam mengumpulkan informasi atau data yang akan digunakan, kemudian dianalisa dan diklasifikasi untuk memecahkan permasalahan yang telah dirumuskan dengan tahapan penelitian yang ada.

- a. Pengumpulan Data Data penelitian diperoleh dari website bpss.go.id
- b. Pengelolaan Data Awal Pengelolaan awal pada penelitian ini akan mencakup semua kegiatan untuk mempersiapkan data sebelum masuk proses pemodelan. Dalam melakukan pengolahan data awal, akan dilakukan beberapa tahapan agar pada akhirnya akan didapatkan data yang bisa digunakan pada tahap berikutnya. Kemudian dilakukan pembersihan data, menghilangkan menghapus data yang tidak lengkap (outer/noise) untuk

memberikan hasil data yang berkualitas dan informatif. Tahapan tersebut antara lain: cleaning, reduction, transformation serta akan dilakukan split validation.

- c. Metode yang Diusulkan Pada tahapan ini dijelaskan pemilihan dan penggunaan metode algoritma C.45.
- d. Pengujian/Validasi Hasil Memberikan hasil penentuan warga mampu dan tidak mampu dengan perhitungan atribut yang digunakan. (Muhidinetal.,2019).

4. HASIL DAN PEMBAHASAN

4.1. Data Set

Pengolahan data adalah proses manipulasi dan transformasi data untuk memperoleh informasi yang relevan. Ini mencakup tahapan pembersihan, analisis, dan interpretasi data untuk membuat keputusan berbasis informasi. Tabel 1. merupakan sample dataset yang digunakan pada penelitian ini.

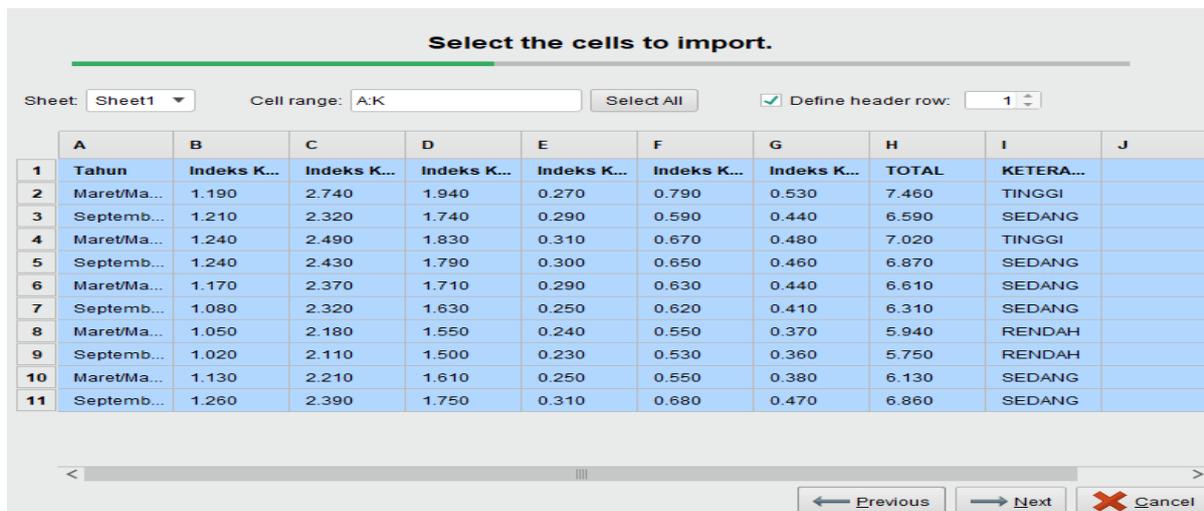
Tabel 1. Data set

Propinsi	Jumlah Penduduk Miskin (000)			Persentase Penduduk Miskin (%)			Garis Kemiskinan (Rp)		
	Kota	Desa	Kota+Desa	Kota	Desa	Kota+Desa	Kota	Desa	Kota+Desa
Aceh	165,4	711,1	876,6	12,47	20,97	18,58	352056	310089	321893
Sumatera Utara	669,4	709,1	1 378,40	10,28	10,53	10,41	295080	249165	271738
Sumatera Barat	124,3	273,6	397,9	6,45	8,99	8	321128	273655	292052
Riau	156,4	324,9	481,3	6,68	8,94	8,05	333933	295582	310603
Jambi	105,3	164,7	270,1	10,53	7,29	8,28	328504	248812	273267
Sumatera Selatan	367,6	674,4	1 042,00	13,29	13,58	13,48	296933	238901	259668
Bengkulu	92,7	217,8	310,5	16,89	17,8	17,51	318881	267273	283252
Lampung	237,9	981,1	1 219,00	11,88	16,96	15,65	297421	251202	263088
Kepulauan Bangka Belitung	24	46,2	70,2	3,73	6,96	5,37	374284	390294	382412
Kepulauan Riau	106,6	24,6	131,2	6,77	7,08	6,83	373725	316963	363450
DKI Jakarta	366,8	-	366,8	3,7	0	3,7	392571		392571
Jawa Barat	2 560,00	1 861,50	4 421,50	8,71	12,13	9,89	249170	228577	242104
Jawa Tengah	1 946,50	2 916,90	4 863,40	13,11	16,55	14,98	245817	223622	233769
DI Yogyakarta	306,5	255,6	562,1	13,1	21,29	15,88	284549	241975	270110
Jawa Timur	1 606,00	3 354,60	4 960,50	8,9	16,88	13,08	253947	234556	243783
Banten	333,5	314,8	648,3	4,41	8,31	5,71	262371	228794	251161
Bali	93,2	67,7	161	3,81	4,17	3,95	270020	230389	254221

4.2. Pengolahan Data

Dengan penerapan model C4.5 pada data baru, hasil analisis dapat digunakan untuk mengklasifikasikan tingkat risiko dengan lebih efisien. Proses ini menyediakan gambaran prediktif berdasarkan pola yang teridentifikasi dalam dataset

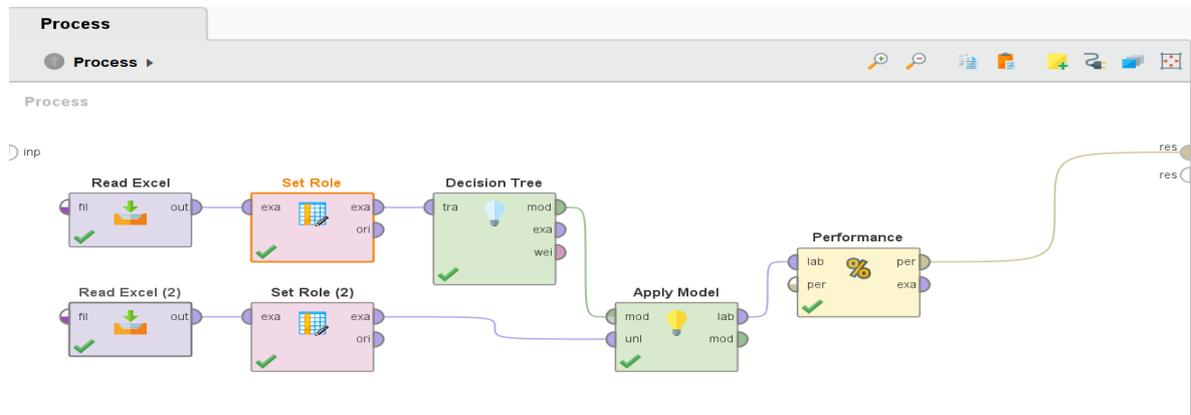
pelatihan, sehingga mendukung pengambilan keputusan yang berbasis pada data yang telah diproses. Berikut adalah tampilan data yang telah dimasukkan ke dalam aplikasi RapidMiner:



Gambar 2. Data Olah Dalam Aplikasi RapidMiner

Tujuan utama dari tugas deskriptif adalah mengenali pola-pola seperti korelasi, tren, kluster, trajektori, dan anomali yang merangkum hubungan utama dalam data. Tugas data mining deskriptif umumnya bersifat eksploratif dan sering

memerlukan proses pascapemrosesan untuk memvalidasi serta memberikan penjelasan atas hasil yang didapatkan. Berikut ini adalah tampilan desain performa menggunakan Algoritma C4.5.



Gambar 3. Desain Performa Menggunakan Algoritma C4.5

accuracy: 5.88%

	true Kota	true 12.47	true 10.28	true 6.45	true 6.68	true 10.53	true 13.29	true 16.89	true 11.88	true 3.73	true 6.77
pred. Kota	1	0	0	0	0	0	0	0	0	0	0
pred. 12....	0	0	0	0	0	0	0	0	0	0	0
pred. 10....	0	0	0	0	0	0	0	0	0	0	0
pred. 6.45	0	0	0	0	0	0	0	0	0	0	0
pred. 6.68	0	0	0	0	0	0	0	0	0	0	0
pred. 10....	0	0	0	0	0	0	0	0	0	0	0
pred. 13....	0	0	0	0	0	0	0	0	0	0	0
pred. 16....	0	0	0	0	0	0	0	0	0	0	0
pred. 11....	0	0	0	0	0	0	0	0	0	0	0
pred. 3.73	0	0	0	0	0	0	0	0	0	0	0
pred. 6.77	0	0	0	0	0	0	0	0	0	0	0
pred. 3.7	0	0	0	0	0	0	0	0	0	0	0
pred. 8.7...	0	0	0	0	0	0	0	0	0	0	0
pred. 13....	0	0	0	0	0	0	0	0	0	0	0

Gambar 4. Data Performance

Data pada gambar 4 digunakan untuk mengevaluasi kinerja model klasifikasi. Confusion matrix ini digunakan untuk menganalisis kemampuan model dalam mengklasifikasikan data ke dalam beberapa kategori atau kelas. Berikut adalah penjelasan terkait data tersebut:

a. Performance Vector:

accuracy: 5.88%: Ini menunjukkan bahwa model memiliki tingkat akurasi yang sangat rendah, hanya 5.88%. Artinya, model tersebut hanya benar dalam memprediksi kelas sekitar 5.88% dari total data yang diuji.

b. Confusion Matrix: Ini adalah tabel yang menunjukkan bagaimana model memprediksi setiap kelas dibandingkan dengan nilai yang sebenarnya. Biasanya, baris pada *confusion matrix* mewakili kelas sebenarnya, dan kolom mewakili kelas yang diprediksi oleh model.

- Kelas yang tercantum di bagian atas (seperti "Kota", "12.47", "10.28", dll.) mewakili kelas yang diprediksi oleh model.

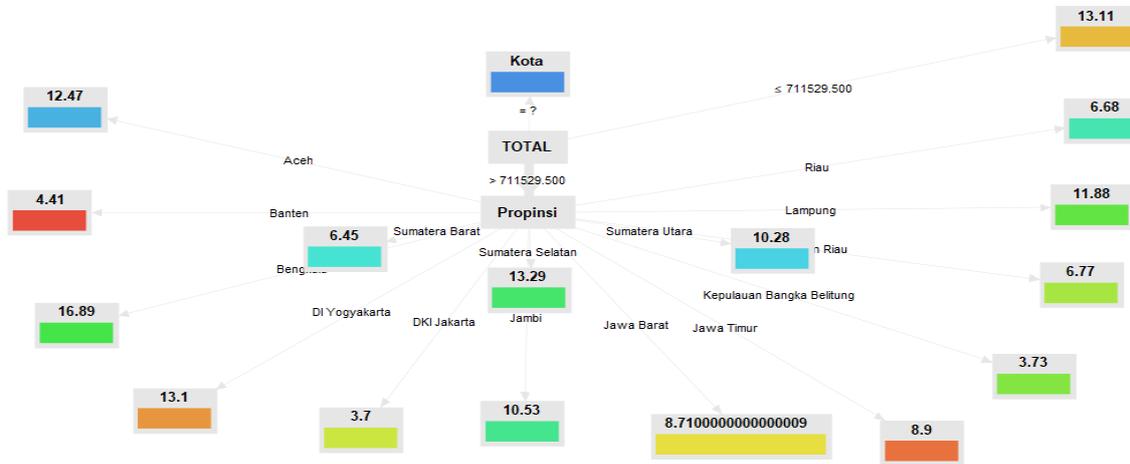
- Nilai-nilai dalam tabel (seperti 1, 0, 0, dll.) menunjukkan jumlah prediksi yang sesuai atau tidak sesuai untuk setiap kombinasi kelas yang sebenarnya (disebut sebagai baris) dan kelas yang diprediksi (disebut sebagai kolom).
- c. Interpretasi Matrix: Baris pertama yang tertera adalah *True: Kota*, yang berarti bahwa data yang benar-benar termasuk dalam kelas "Kota" ada pada baris pertama.
- Dalam kolom pertama, nilai 1 menunjukkan bahwa model memprediksi kelas "Kota" dengan benar 1 kali.
 - Pada kolom kedua sampai kolom terakhir, nilai-nilai adalah 0, yang menunjukkan bahwa model tidak membuat prediksi yang benar untuk kelas lainnya.
- d. Kelas yang Tercatat: Kelas-kelas seperti "12.47", "10.28", dan seterusnya mungkin adalah label numerik atau nama kategori yang digunakan dalam sistem Anda. Namun, tanpa konteks lebih lanjut, ini tidak dapat dijelaskan lebih detail.

e. Kesimpulan Kinerja:

- Model memiliki akurasi yang sangat rendah, hanya 5.88%. Ini berarti model tersebut tidak dapat mengklasifikasikan data dengan benar dalam sebagian besar kasus.
- Semua nilai dalam confusion matrix untuk kombinasi kelas yang berbeda tampaknya adalah 0, menunjukkan bahwa model hampir

tidak pernah membuat prediksi yang benar untuk kelas lainnya (selain "Kota").

Secara umum, model ini membutuhkan peningkatan atau pelatihan tambahan untuk meningkatkan tingkat akurasinya. Berdasarkan data di atas, dapat dibuat pohon keputusan seperti yang ditunjukkan pada Gambar 5.



Gambar 5. Pohon Biner

Pohon biner merupakan representasi dari suatu model klasifikasi yang menggunakan pembagian berdasarkan nilai tertentu untuk memprediksi kategori tertentu (seperti propinsi atau daerah). Berikut adalah penjelasan mengenai struktur pohon biner yang tertera:

Struktur Pohon Biner:

a. Total

- Ini merupakan kondisi yang menggambarkan total data yang akan diprediksi atau diobservasi, dengan nilai yang lebih besar dari 711529.500.
- Pada titik ini, pohon biner tidak memberikan informasi lebih lanjut mengenai apa yang dilakukan dengan data ini, selain memulai pembagian berdasarkan nilai yang lebih besar daripada threshold tersebut.

b. Pembagian Berdasarkan Propinsi dan Nilai:

- Di bawah TOTAL > 711529.500, setiap cabang pohon memeriksa kondisi lebih lanjut berdasarkan nilai dari suatu propinsi tertentu. Setiap cabang memiliki kondisi yang memeriksa propinsi (seperti Aceh, Banten, Bengkulu, dll.).
- Misalnya:
 - Propinsi = Aceh: Nilai yang diprediksi adalah 12.47. Jika propinsi yang sedang dianalisis adalah Aceh, maka kategori yang diprediksi adalah 12.47, dan semua kategori lain (seperti Kota, 10.28, 6.45, dll.) akan diprediksi sebagai 0 (tidak terjadi).

- Propinsi = Banten: Nilai yang diprediksi adalah 4.41, dan kategori yang diprediksi hanya 4.41 yang mendapat nilai 1, sementara kategori lainnya 0.
- Propinsi = Bengkulu: Nilai yang diprediksi adalah 16.89, dengan kategori 16.89 yang diprediksi sebagai 1, dan kategori lainnya 0.

c. Kondisi Lain (≤ 711529.500):

- Setelah memeriksa apakah data lebih besar dari 711529.500, cabang lainnya menunjukkan kondisi di mana nilai totalnya lebih kecil atau sama dengan 711529.500, yang mana propinsi yang diprediksi adalah 13.11.
- Di bawah kondisi ini, hanya kategori 13.11 yang diprediksi sebagai 1, dan kategori lainnya tetap 0.

d. Penjelasan Logika Pengambilan Keputusan:

Pohon ini membagi data berdasarkan nilai-nilai tertentu yang terkait dengan propinsi di Indonesia, dan untuk setiap propinsi, model mengklasifikasikan data ke dalam kategori yang berbeda.

- Setiap cabang (misalnya Aceh, Banten, dll.) mewakili keputusan berdasarkan lokasi atau propinsi, dan hasilnya adalah angka yang mengindikasikan kategori mana yang diprediksi oleh model.
- Misalnya, jika data berasal dari Aceh, maka nilai 12.47 adalah hasil prediksi model, dan semua kategori lain diberikan nilai 0.

- Begitu juga dengan propinsi lain seperti Banten, Bengkulu, dan sebagainya, masing-masing memiliki angka tertentu yang diprediksi oleh model.
- Pohon ini menggambarkan pembagian keputusan klasifikasi yang bergantung pada propinsi dan nilai tertentu yang terkait dengan setiap propinsi tersebut.
- Setiap cabang pohon menunjukkan keputusan berdasarkan apakah propinsi itu cocok dengan kondisi tertentu, dan hasil prediksi yang berbeda akan diberikan sesuai dengan nilai yang tercatat untuk propinsi tersebut.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil perancangan, implementasi, dan pengujian yang dilakukan, dapat diambil kesimpulan bahwa Implementasi algoritma C4.5 untuk prediksi Indeks Kemiskinan Daerah dimulai dengan melakukan analisis kebutuhan data kemudian dilakukan penelitian, setelah didapatkan data-data yang diperlukan kemudian dilakukan pembuatan sistem dengan mengimplementasikan algoritma C4.5 kedalam sistem tersebut dan melakukan pengujian terhadap beberapa sampel data untuk mendapatkan akurasi dari hasil prediksi yang dilakukan oleh sistem yang telah dibuat. Penelitian ini menunjukkan bahwa kemiskinan di Indonesia, meskipun mengalami penurunan dalam dimensi moneter, masih menjadi tantangan besar dalam dimensi multidimensi yang melibatkan kesehatan, pendidikan, dan standar hidup. Wilayah dengan ketimpangan pembangunan, terutama di kawasan timur Indonesia, menghadapi tingkat kemiskinan multidimensi yang lebih tinggi dibandingkan wilayah lain. Faktor utama yang berkontribusi terhadap kemiskinan ini meliputi keterbatasan akses terhadap layanan kesehatan, rendahnya kualitas pendidikan, dan kurangnya infrastruktur dasar. Untuk mengurangi kemiskinan secara efektif, pendekatan multidimensi harus menjadi landasan utama dalam perumusan kebijakan. Rekomendasi mencakup penguatan program perlindungan sosial, peningkatan akses layanan dasar, serta pemerataan pembangunan infrastruktur di seluruh wilayah Indonesia. Implementasi strategi yang inklusif dan berkelanjutan akan mempercepat upaya pengentasan kemiskinan dan mendorong tercapainya tujuan pembangunan berkelanjutan di Indonesia. Temuan ini memberikan wawasan penting bagi para pemangku kebijakan dalam menyusun langkah strategis untuk menghadapi tantangan kemiskinan di masa depan. Indeks keparahan kemiskinan cenderung memiliki tren yang berbeda, meskipun pernah terjadi kenaikan yang cukup tajam pada tahun 2021. Jika dilihat menurut daerah, indeks keparahan kemiskinan di perkotaan terlihat lebih berfluktuasi.

DAFTAR PUSTAKA

- [1] E. Adawiyah, "Kemiskinan Dan Faktor-Faktor Penyebabnya," vol. 1, no. April, pp. 43–50, 2020.
- [2] C. Maharani, D. A. Ningrum, A. E. Fatmawati, and A. Fadilla, "Dampak Kemiskinan terhadap Kualitas Pendidikan Anak di Indonesia: Rekomendasi Kebijakan yang Efektif," *J. Macroecon. Soc.*, vol. 1, no. 3, pp. 1–10, 2024, doi: 10.47134/jmsd.v1i3.199.
- [3] P. Desa and M. Sari, "Penerapan metode c4.5 untuk klasifikasi warga miskin pada desa mengandung sari," vol. 2, no. 3, pp. 14–25, 2021.
- [4] L. N. Aini and S. N. Islamy, "Dampak pengangguran, pendidikan, kesehatan, PDRB dan indeks pembangunan manusia terhadap kemiskinan di Indonesia," *J. Econ. Res. Policy Stud.*, vol. 1, no. 3, pp. 132–141, 2021, doi: 10.53088/jerps.v1i3.325.
- [5] S. Frisnoiry, I. E. Putri, N. I. Sari, and C. R. A. K. Sinaga, "Analisis Kemiskinan Terhadap Ketimpangan Pendapatan di Indonesia pada Tahun 2023," *EKOMA J. Ekon. Manajemen, Akunt.*, vol. 3, no. 5, pp. 1360–1367, 2024.
- [6] A. Muhidin, P. Studi, T. Informatika, S. Tinggi, and T. Pelita, "Volume 9 Nomor 3 Maret 2019 ISSN : 2407-3903," vol. 9, pp. 13–18, 2019.
- [7] S. D. Purnomo and K. S. Danuta, "Analisis Kemampuan Keuangan Daerah terhadap Kemiskinan: Studi Empiris di Sumatera Utara," *Ekon. J. Econ. Bus.*, vol. 6, no. 1, p. 215, 2022, doi: 10.33087/ekonomis.v6i1.513.
- [8] R. P. P. Sinurat, "Analisis Faktor-Faktor Penyebab Kemiskinan Sebagai Upaya Penanggulangan Kemiskinan Di Indonesia," *J. Regist.*, vol. 5, no. 2, pp. 87–103, 2023, doi: 10.33701/jurnalregistratie.v5i2.3554.
- [9] Samsul Hadi, Ahmad Taufiki, Maulana Zakaria Ahmad, and Adil Siswanto, "Analisis Tingkat Kemiskinan Tahun 2023 Di Badan Pusat Statistik (BPS) Kabupaten Jember," *J. Indones. Soc. Soc.*, vol. 2, no. 1, pp. 32–39, 2024, doi: 10.59435/jiss.v2i1.224.
- [10] A. Maulani Nisa, Rusdarti, and A. Wahyudin, "Pengaruh Pertumbuhan Ekonomi, Investasi dan Belanja Publik Terhadap Kemiskinan dengan Indeks Pembangunan Manusia (IPM) sebagai Variabel Moderasi di Jawa Tengah," *J. Ekon. Dan Bisnis*, vol. 3, no. 2, pp. 65–74, 2023.
- [11] C. I. Ifaru *et al.*, "Penentuan Sparepart Genset Paling Sering Digunakan Pada Operator Indosat Ooredoo Hutchison Dengan Algoritma C4.5," *J. Teknol. Sist. Inf.*, vol. 5, no. 1, pp. 300–309, 2024, doi: 10.35957/jtsi.v5i1.7481.
- [12] P. Riliandhita, Iqbal Maulana, "Klasifikasi Penentuan Status Kemiskinan Penduduk Kelurahan," vol. 8, no. 2, pp. 1791–1796, 2024.
- [13] D. M. Musa *et al.*, "Penerapan Data Mining Untuk Klasifikasi Data Penjualan Pakan Ternak

- Terlaris Dengan Algoritma C4.5,” *J. Teknol. Inform. dan Komput.*, vol. 10, no. 1, pp. 168–182, 2024, doi: 10.37012/jtik.v10i1.1985.
- [14] M. A. Hidayatulloh, K. P. Kartika, and D. F. H. Permadi, “Implementasi Algoritma C4.5 Untuk Memprediksi Capaian Pembelajaran Daring (Studi Kasus Siswa MAN 3 Blitar),” *J. Algoritm.*, vol. 3, no. 1, pp. 33–47, 2022, doi: 10.35957/algoritme.v3i1.3292.
- [15] D. Jatmika and G. Taufik, “Penerapan Algoritma C4.5 Untuk Klasifikasi Keberhasilan Pengiriman Barang,” *Invotek Polbeng*, vol. 8, no. 1, pp. 12–26, 2021.
- [16] S. Faisal, “Klasifikasi Data Mining Menggunakan Algoritma C4.5 Terhadap Kepuasan Pelanggan Sewa Kamera Cikarang,” *Techno Xplore J. Ilmu Komput. dan Teknol. Inf.*, vol. 4, no. 1, pp. 1–8, 2019, doi: 10.36805/technoexplore.v4i1.541.
- [17] S. Setiawansyah, “Penerapan Metode Entropy dan Grey Relational Analysis dalam Evaluasi Kinerja Karyawan,” *J. Data Sci. Inf. Syst.*, vol. 2, no. 1, pp. 29–39, 2024.