

IMPLEMENTASI ALGORITMA C4.5 PADA VARIABEL-VARIABEL YANG MEMPENGARUHI STATUS GIZI BALITA DESA KUALA DUA KABUPATEN KUBU RAYA

Supriyati, Hendra Perdana

INTISARI

Permasalahan gizi sering terjadi pada anak usia dibawah lima tahun. Masa pertumbuhan pada anak usia balita menjadi perhatian serius para orang tua. Banyak faktor yang mempengaruhi status gizi balita, antara lain keluarga, sarana dan prasarana kesehatan, dan keadaan lingkungan sekitar. Keberagaman faktor-faktor tersebut dapat diklasifikasikan dengan menggunakan metode pohon keputusan. Pohon keputusan adalah suatu metode klasifikasi serta prediksi yang berguna untuk mengeksplorasi data dan menemukan hubungan antar variabel. Ada beberapa algoritma yang dapat digunakan dalam pembentukan pohon keputusan, diantaranya adalah Algoritma C4.5. Dalam penelitian ini, Algoritma C4.5 digunakan untuk menganalisis dan mengetahui variabel-variabel yang mempengaruhi status gizi balita Desa Kuala Dua Kabupaten Kubu Raya. Penilaian status gizi anak dihitung berdasarkan standar pengukuran status gizi. Perhitungan dimulai dengan menghitung nilai entrophy total, entrophy masing-masing atribut dan gain masing-masing atribut. Pohon keputusan yang terbentuk menjadikan variabel sumber air yang digunakan sebagai akar pohon karena memiliki nilai gain terbesar yaitu 0,3330. Pohon keputusan terdiri dari 2 tingkatan yang saling berhubungan antar variabel-variabel yang mempengaruhi status gizi balita.

Kata Kunci: C4.5, gizi, gain, entrophy

PENDAHULUAN

Kualitas sumber daya manusia dipengaruhi oleh aspek kesehatan, aspek pendidikan dan aspek pendapatan. Aspek kesehatan mempunyai peranan penting dalam meningkatkan kesejahteraan hidup masyarakat. Definisi sehat menurut WHO adalah keadaan tubuh ideal dari aspek fisik, mental, sosial dan terbebas dari berbagai penyakit. Perubahan gaya hidup dan pola makan membuat masyarakat mengalami permasalahan gizi terutama balita (bayi usia dibawah lima tahun). Usia balita merupakan usia yang masih bergantung pada peran orangtua dalam hal pengasuhan maupun asupan makanan [1]. Permasalahan gizi disebabkan oleh faktor penyebab langsung dan faktor penyebab tidak langsung. Faktor penyebab langsung yaitu makanan anak dan penyakit yang pernah diderita sedangkan faktor penyebab tidak langsung yaitu berupa ketahanan pangan, pola pengasuhan, pelayanan kesehatan masyarakat dan kesehatan lingkungan masyarakat [2].

Kabupaten Kubu Raya (KKR) merupakan kabupaten pemekaran dari Kabupaten Pontianak (sekarang Kabupaten Mempawah). Secara keseluruhan, wilayah administratif KKR terbagi dalam 9 kecamatan. Jumlah populasi tertinggi KKR berada di wilayah Kecamatan Sungai Raya sebanyak 198,885 jiwa atau sekitar 38 % dari total seluruh populasi KKR sebesar 529,320 jiwa [3]. Tingginya populasi yang berada di daerah Kecamatan Sungai Raya tentunya membuat daerah ini perlu mendapat perhatian serius dari berbagai aspek terutama aspek kesehatan. Salah satu desa yang berada di Kecamatan Sungai Raya antara lain Desa Kuala Dua. Desa ini memiliki jumlah balita tertinggi laki-laki dan perempuan sebanyak 2.441 jiwa dari total 9.488 jiwa. Ditinjau dari status gizi balita Tahun 2012, ditemukan sebanyak 401 responden balita balita berstatus gizi kurang atau 9,10 % dan 94 responden balita balita berstatus gizi buruk atau 2,13% dari 7.497 jumlah balita [4].

Beberapa faktor yang melatarbelakangi kasus-kasus gizi tersebut antara lain, di Desa Kuala Dua partisipasi masyarakat untuk menimbangkan anaknya ke posyandu masih tergolong rendah, dilihat dari persentasi penimbangan balita hanya 47,73% balita yang ditimbang dari 4,408 balita yang ada, dan dari seluruh penimbangan tersebut sebanyak 74,82% balita naik berat badannya. Kurangnya partisipasi masyarakat untuk menimbangkan balitanya secara rutin ke posyandu dikarenakan

kurangnya pengetahuan tentang fungsi posyandu. Posyandu selama ini hanya dianggap sebagai tempat untuk imunisasi balita dan distribusi vitamin A. Selain itu dari sisi pola asuh, kurangnya pengetahuan ibu tentang ASI eksklusif, Makanan Pendamping ASI (MP ASI) yang kurang tepat waktu pemberiannya dan minimnya pengetahuan ibu terhadap pemilihan bahan makanan yang bergizi selain juga faktor ekonomi (daya beli), semakin mempengaruhi kecenderungan balita mengalami kurang gizi yang dapat mengakibatkan terjadinya gizi buruk [4].

Keberagaman faktor-faktor penyebab tersebut dapat diklasifikasikan menggunakan suatu metode pohon keputusan. Pohon keputusan merupakan suatu metode klasifikasi serta prediksi yang berguna untuk mengeksplorasi data dan menemukan hubungan antar variabel. Sebuah model pohon keputusan terdiri dari sekumpulan aturan untuk membagi sejumlah populasi yang heterogen. Salah satu metode pohon keputusan adalah Algoritma C4.5 (*Classification 4.5 Algorithm*) [5].

Secara umum, sebuah pohon dibentuk oleh akar, daun, dan cabang. Proses klasifikasi menempatkan atribut dengan nilai *gain* tertinggi sebagai akar sedangkan pembentukan simpul ditentukan melalui nilai *entropy* terbesar yang dimiliki masing-masing atribut. Nilai atribut merupakan nilai-nilai hasil pengamatan, nilai *entropy* merupakan nilai rata-rata dari masing-masing kelompok data sedangkan nilai *gain* merupakan nilai yang menyatakan pengaruh suatu atribut terhadap kasus yang diteliti.

Dalam penelitian ini, penulis menerapkan Algoritma C4.5 untuk melihat variabel-variabel yang mempengaruhi status gizi balita. Penelitian ini bertujuan untuk membentuk suatu pohon keputusan serta menjelaskan keterkaitan antar variabel-variabel yang mempengaruhi status gizi balita Desa Kuala Dua Kabupaten Kubu Raya. Adapun variabel-variabel yang menjadi bahan penelitian ini meliputi pendidikan terakhir ibu balita, pekerjaan ayah balita, jumlah anak yang dimiliki dalam satu keluarga balita, penyakit serius yang pernah diderita balita, pemberian ASI, lama pemberian ASI, pemberian imunisasi, intensitas makan balita dalam sehari, porsi makan balita dalam sehari, sumber air bersih yang digunakan keluarga, tempat menyimpan makanan balita serta pengolahan makanan.

Adapun variabel-variabel tersebut bersumber pada data primer yang penulis peroleh dari kuesioner hasil wawancara terhadap kondisi keluarga, pola asuh serta keadaan lingkungan sekitar masyarakat Desa Kuala Dua. Penelitian dilakukan selama tiga hari di Dusun Keramat I, Dusun Keramat II, dan Dusun Karya I. Populasi dalam penelitian ini yaitu balita dan keluarga balita yang tercatat pada laporan posyandu bulan Januari Tahun 2016 di Dusun Keramat I, Dusun Keramat II, dan Dusun Karya I. Sampel dalam penelitian ini sebanyak 30 sampel dan cara pengambilan sampel menggunakan teknik *probability sampling*.

Sebelum menerapkan Algoritma C4.5 terlebih dahulu menentukan status gizi balita. Data yang digunakan dalam penentuan status gizi balita yaitu data sekunder laporan posyandu bulan Januari Tahun 2016 di Dusun Keramat I, Dusun Keramat II, dan Dusun Karya I berupa data berat badan, tinggi badan, umur dan jenis kelamin balita. Penentuan status gizi balita diperoleh dengan menghitung nilai *Z*. Nilai *Z* yang diperoleh selanjutnya diklasifikasikan kedalam kategori standar status gizi pada tabel *antropometri* yang dikeluarkan oleh Menkes RI, Tahun 2011 dalam Keputusan Nomor : 1995/MENKES/SK/XII/2010 tentang standar *antropometri*. Dalam penelitian ini, indeks *antropometri* yang digunakan adalah berat badan/umur (BB/U).

Setelah menentukan status gizi, langkah selanjutnya yaitu menerapkan variabel-variabel penelitian kedalam Algoritma C4.5. Penelitian ini menggunakan tiga kasus status gizi yaitu gizi baik, gizi kurang dan gizi buruk. Proses perhitungan pada Algoritma C4.5 ini diawali dengan menghitung nilai *entropy* total, nilai *entropy* masing-masing atribut ke-*i* dan nilai *gain* masing-masing atribut. Perhitungan nilai *entropy* dan nilai *gain* dilakukan dengan bantuan *Microsoft Excel*. Nilai *gain* terbesar yang didapat dijadikan sebagai simpul akar (*root*) dalam membentuk sebuah pohon keputusan. Kemudian nilai *entropy* digunakan dalam membentuk cabang pohon keputusan. Perhitungan selesai apabila atribut telah menghasilkan sebuah nilai solusi dari tiga kasus yang digunakan (gizi baik, gizi kurang dan gizi buruk).

STATUS GIZI

Kualitas sumber daya manusia (SDM) dipengaruhi oleh tiga faktor yaitu kesehatan, pendidikan dan keadaan ekonomi tiap individu. Faktor kesehatan dapat dilihat melalui status gizi. Status gizi merupakan keadaan tubuh akibat dari konsumsi makanan serta faktor-faktor yang mempengaruhi kehidupan tiap individu. Upaya peningkatan kualitas SDM dimulai dari penanganan pertumbuhan balita sebagai bagian dari keluarga. Usia balita merupakan usia yang masih bergantung pada peran orangtua dalam hal asupan gizi dan pengasuhan [1].

Status gizi balita dipengaruhi oleh faktor penyebab langsung dan faktor penyebab tidak langsung. Faktor penyebab langsung adalah konsumsi makanan dan penyakit yang pernah diderita. Sedangkan faktor penyebab tidak langsung permasalahan gizi yaitu ketahanan pangan keluarga, pola pengasuhan anak serta pelayanan kesehatan dan kesehatan lingkungan. Adapun faktor penyebab tidak langsung dari sisi lain yaitu berkaitan dengan tingkat pendidikan, pengetahuan, dan keterampilan keluarga [2].

Menteri Kesehatan Republik Indonesia (Menkes RI) telah menetapkan standar penilaian status gizi yang dikenal dengan *antropometri*. Istilah *antropometri* menurut bahasa artinya ukuran tubuh sehingga indeks yang digunakan berupa berat badan, tinggi badan, dan indeks massa tubuh. Dalam Keputusan Menkes RI Nomor:1995/MENKES/SK/XII/2010, menggunakan tiga indeks dalam mengukur status gizi anak yaitu berat badan/umur (BB/U), tinggi badan/umur (TB/U), berat badan/tinggi badan (BB/TB) dan indeks massa tubuh/umur (IMT/U). Dalam penelitian ini menggunakan indeks berat badan/umur (BB/U). Penentuan status gizi didasarkan pada perhitungan nilai Z yang diperoleh dari perhitungan nilai subyek individu dikurangi nilai median baku rujukan dibagi dengan nilai standar deviasi (SD) atau simpangan baku rujukan. Dalam tabel *antropometri*, nilai median yang dimaksud adalah nilai yang berada pada kolom nilai median yang sejajar dengan nilai pada baris usia. Sedangkan nilai standar deviasi yang dimaksud merupakan nilai yang berada pada kolom nilai standar deviasi yang sejajar dengan nilai pada baris usia. Hasil dari perhitungan nilai Z kemudian disesuaikan dengan kategori status gizi yang telah ditetapkan. Secara matematis rumus untuk menghitung nilai Z sebagai berikut:

$$Z = \frac{\text{Nilai subyek penelitian} - \text{Nilai median rujukan}}{\text{Nilai standar deviasi rujukan}}$$

Perolehan nilai Z dikonversi dengan nilai SD yang terdapat pada tabel *antropometri*. Setelah itu status gizi anak dapat dikelompokkan kedalam kategori-kategori yang telah ditentukan [6]. Kategori penilaian status gizi dapat dilihat pada Tabel 1.

Tabel 1 Kategori Penilaian Status Gizi berdasarkan Indeks BB/U

Indeks	Batas Nilai Z_{score}	Status Gizi
BB/U (0-60 Bulan)	< -3	Gizi buruk
	- 3 s/d < -2	Gizi kurang
	- 2 s/d 2	Gizi baik
	> 2	Gizi lebih

Sumber : Kemenkes RI, 2011

POHON

Konsep pohon pertama kali digunakan pada Tahun 1857 oleh seorang matematikawan Inggris yang bernama Arthur Cayley untuk menghitung senyawa. Graf dan pohon merupakan suatu konsep yang saling berhubungan. Pohon merupakan bagian dari graf dan menggunakan konsep yang diterapkan oleh graf. Pengertian pohon secara umum merupakan sebuah graf terhubung yang tidak memuat sirkuit. Suatu graf dikatakan terhubung, apabila setiap pasang simpul didalam graf terhubung. Sedangkan sirkuit merupakan lintasan tertutup dengan simpul awal dan simpul akhir sama. Pengertian

pohon keputusan merupakan serangkaian keputusan yang menghasilkan sebuah solusi. Tiap simpul menyatakan keputusan sedangkan daun menyatakan sebuah solusi dari permasalahan. Sebuah pohon terdiri dari akar, cabang, dan simpul yang akan berkaitan [7].

Istilah-istilah yang sering digunakan dalam pohon adalah akar (*root*), cabang, daun (*leaf*), tingkat atau kedalaman (*depth*), dan derajat (*degree*). Akar (*root*) simpul awal pembentukan pohon dan dijadikan sebagai dasar penentuan simpul-simpul selanjutnya. Selain akar, bagian yang berperan penting dalam pembentukan sebuah pohon adalah cabang. Cabang merupakan bagian yang menghubungkan antara simpul-simpul lainnya. Kemudian daun (*leaf*) merupakan simpul yang berada dibawah simpul akar dan tidak terdapat cabang pembentuk simpul selanjutnya. Tingkat atau kedalam (*depth*) diartikan sebagai suatu keadaan dimana setiap simpul turun satu tingkat membentuk sebuah simpul baru [8].

ALGORITMA C4.5

Algoritma C4.5 merupakan salah satu Algoritma dalam pohon keputusan. Proses klasifikasi mengelompokkan nilai atribut kedalam kasus. Atribut merupakan variabel-variabel hasil pengamatan. Adapun langkah-langkah pengerjaan dari Algoritma C4.5 sebagai berikut [5]

1. Pembentukan simpul akar berdasarkan nilai atribut dengan nilai *gain* tertinggi
2. Pembentukan cabang tiap nilai berdasarkan nilai *entropy*
3. Pembentukan simpul berhenti setelah simpul menghasilkan solusi dari kasus yang diteliti

Nilai *entropy* merupakan nilai rata-rata dari masing-masing suatu kumpulan data. Adapun rumus untuk menghitung nilai *entropy* adalah:

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

dengan S adalah himpunan seluruh atribut, p_i adalah proporsi p_i terhadap S , n adalah jumlah partisi S .

Ada dua nilai *entropy* yang digunakan yaitu *Entropy S* dan *Entropy S_i*. *Entropy S* merupakan nilai *Entropy* total sedangkan *Entropy S_i* merupakan nilai *entropy* masing-masing atribut. Adapun rumus untuk mendapatkan nilai *Entropy S_i* adalah:

$$Entropy(S_i) = -\sum_{i=1}^n pq_i \cdot \log_2 pq_i \quad (2)$$

dengan S_i adalah himpunan masing-masing atribut, pq_i adalah proporsi pq_i terhadap S_i , S_i adalah atribut ke- i , n adalah jumlah partisi atribut S .

Kedua nilai *entropy* yaitu *Entropy S* dan *Entropy S_i* ini digunakan untuk mencari nilai *gain*. Nilai *gain* tersebut akan digunakan dalam setiap penentuan simpul-simpul. Besarnya nilai *gain* menunjukkan seberapa besar pengaruh suatu atribut terhadap pembentukan simpul. Adapun rumus untuk mendapatkan nilai *Entropy S_i* adalah:

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} \cdot Entropy(S_i) \quad (3)$$

dengan S adalah himpunan seluruh atribut, A adalah atribut, n adalah jumlah partisi S , $|S_i|$ adalah jumlah kasus masing-masing atribut ke- i , $|S|$ adalah jumlah kasus seluruh atribut dalam S .

APLIKASI NUMERIK

Penelitian ini menerapkan Algoritma C4.5 pada variabel-variabel yang mempengaruhi status gizi balita Desa Kuala Dua Kabupaten Kubu Raya. Adapun variabel-variabel dalam penelitian ini dapat dilihat pada Tabel 2.

Tabel 2 Variabel-Variabel yang Mempengaruhi Status Gizi Balita

Variabel-Variabel	
Pendidikan Ibu (A)	SD (A_1) SMP (A_2) SMA (A_3)
Pekerjaan Ayah (B)	Swasta (B_1) Buruh (B_2)
Jumlah Anak Yang Dimiliki Keluarga Balita (C)	< 4 Anak (C_1) > 4 Anak (C_2)
Penyakit Serius Yang Pernah Diderita Balita (D)	Pernah (D_1) Tidak Pernah (D_2)
Pemberian ASI (E)	Ya (E_1) Tidak (E_2)
Lama Pemberian ASI (F)	< 1 Tahun (F_1) > 1 Tahun (F_2)
Pelaksanaan Imunisasi Balita (G)	Rutin (G_1) Tidak Rutin (G_2)
Intensitas Makan Balita Dalam Sehari (H)	2 Kali (H_1) 3 Kali (H_2)
Porsi Makan Balita Dalam Sehari (I)	Cukup (I_1) Kurang (I_2)
Sumber Air Bersih Yang Digunakan (J)	Penampungan Air Hujan (J_1) Sumur (J_2) Sungai/Parit (J_3) PDAM (J_4)
Tempat Menyimpan Makanan (K)	Didalam Lemari (K_1) Diatas Meja (K_2)
Pengolahan Makanan (L)	Direbus (L_1) Direbus dan Digoreng (L_2)

Sebelum mengaplikasikan Algoritma C4.5 terlebih dahulu dilakukan penentuan status gizi balita menggunakan rumus nilai Z kemudian dikonversikan dengan standar *antropometri* yang dikeluarkan oleh Menkes RI Tahun 2010. Indeks *antropometri* yang digunakan adalah berat badan/umur (BB/U). Total 30 responden masing-masing dicari kategori status gizi berdasarkan data berat badan, usia dan jenis kelamin. Hasil dari perhitungan nilai Z dapat dilihat pada Tabel 3.

Tabel 3 Perolehan Nilai Z Masing-Masing Responden Balita

Responden Balita	Nilai Z	Responden balita	Nilai Z
1	-2.53	16	-2.44
2	-2.50	17	-2.64
3	-1.87	18	-2.27
4	-2.75	19	-3.17
5	-0.27	20	-1
6	-1.21	21	-1.92
7	-2.55	22	-1
8	-2.14	23	-2.17
9	-1.27	24	-2.63
10	0.94	25	-1.57
11	-1.82	26	-3.73
12	0.67	27	-1.45
13	-4	28	-2.64
14	-2.11	29	-3.12
15	-2	30	-2.75

Berdasarkan Tabel 3, nilai Z yang berada pada rentang -2 SD s/d 2 SD dengan status gizi baik terdapat 13 responden balita, rentang -3 SD s/d < -2 SD dengan status gizi kurang sebanyak 13 responden balita dan rentang < -3 SD dengan status gizi buruk yaitu 4 responden balita.

Setelah menentukan status gizi masing-masing responden balita maka langkah selanjutnya adalah membentuk pohon keputusan. Pembentukan pohon keputusan diawali dengan membentuk simpul akar. Perhitungan dimulai dengan menghitung nilai *entropy* total, nilai *entropy* masing-masing atribut dan nilai *gain* masing-masing atribut. Nilai *entropy* total akan digunakan pada perhitungan nilai *gain*. Dengan N merupakan total responden balita sebanyak 30 responden balita, $\sum B$ adalah jumlah responden balita berstatus gizi baik sebanyak 13 responden balita, $\sum K$ adalah jumlah responden balita berstatus gizi kurang sebanyak 13 responden balita dan $\sum Bu$ adalah jumlah responden balita berstatus gizi buruk sebanyak 4 responden balita. Diperoleh nilai *entropy* total :

$$\begin{aligned} Entropy_{total} &= \left(\left(-\frac{\sum B}{N} \right) \cdot \log_2 \left(\frac{\sum B}{N} \right) \right) + \left(\left(-\frac{\sum K}{N} \right) \cdot \log_2 \left(\frac{\sum K}{N} \right) \right) + \left(\left(-\frac{\sum Bu}{N} \right) \cdot \log_2 \left(\frac{\sum Bu}{N} \right) \right) \\ &= \left(\left(-\frac{13}{30} \right) \cdot \log_2 \left(\frac{13}{30} \right) \right) + \left(\left(-\frac{13}{30} \right) \cdot \log_2 \left(\frac{13}{30} \right) \right) + \left(\left(-\frac{4}{30} \right) \cdot \log_2 \left(\frac{4}{30} \right) \right) \\ &= 1,4332 \end{aligned}$$

Langkah selanjutnya yaitu menghitung nilai *entropy* masing-masing atribut ke- i . Berdasarkan informasi pada Tabel 2 maka setiap atribut ke- i tersebut dihitung nilai *entrophynya*. Dimulai dengan atribut A_1 dengan $\sum A_1$ adalah jumlah responden balita pada atribut A_1 sebanyak 5 responden balita. Sebagaimana $\sum B_{A_1}$ adalah jumlah responden balita atribut A_1 berstatus gizi baik sebanyak 3 responden balita, $\sum K_{A_1}$ adalah jumlah responden balita atribut A_1 berstatus gizi kurang sebanyak 2 responden balita, dan $\sum Bu_{A_1}$ adalah jumlah responden balita atribut A_1 berstatus gizi buruk berdasarkan perhitungan tidak ada responden balita.

Diperoleh nilai *Entropy* A_1 :

$$\begin{aligned} Entropy(A_1) &= \left(\left(-\frac{\sum B_{A_1}}{\sum A_1} \right) \cdot \log_2 \left(\frac{\sum B_{A_1}}{\sum A_1} \right) \right) + \left(\left(-\frac{\sum K_{A_1}}{\sum A_1} \right) \cdot \log_2 \left(\frac{\sum K_{A_1}}{\sum A_1} \right) \right) + \left(\left(-\frac{\sum Bu_{A_1}}{\sum A_1} \right) \cdot \log_2 \left(\frac{\sum Bu_{A_1}}{\sum A_1} \right) \right) \\ &= \left(\left(-\frac{3}{5} \right) \cdot \log_2 \left(\frac{3}{5} \right) \right) + \left(\left(-\frac{2}{5} \right) \cdot \log_2 \left(\frac{2}{5} \right) \right) + \left(\left(-\frac{0}{5} \right) \cdot \log_2 \left(\frac{0}{5} \right) \right) \\ &= 0,971 \end{aligned}$$

Dengan $\sum A_2$ adalah jumlah responden balita pada atribut A_2 sebanyak 8 responden balita, $\sum B_{A_2}$ adalah jumlah responden balita atribut A_2 berstatus gizi baik sebanyak 4 responden balita, $\sum K_{A_2}$ adalah jumlah responden balita atribut A_2 berstatus gizi kurang sebanyak 3 responden balita, dan $\sum Bu_{A_2}$ adalah jumlah responden balita atribut A_2 berstatus gizi buruk sebanyak 1 responden balita.

Diperoleh nilai *Entropy* A_2 :

$$\begin{aligned} Entropy(A_2) &= \left(\left(-\frac{\sum B_{A_2}}{\sum A_2} \right) \cdot \log_2 \left(\frac{\sum B_{A_2}}{\sum A_2} \right) \right) + \left(\left(-\frac{\sum K_{A_2}}{\sum A_2} \right) \cdot \log_2 \left(\frac{\sum K_{A_2}}{\sum A_2} \right) \right) + \left(\left(-\frac{\sum Bu_{A_2}}{\sum A_2} \right) \cdot \log_2 \left(\frac{\sum Bu_{A_2}}{\sum A_2} \right) \right) \\ &= \left(\left(-\frac{4}{8} \right) \cdot \log_2 \left(\frac{4}{8} \right) \right) + \left(\left(-\frac{3}{8} \right) \cdot \log_2 \left(\frac{3}{8} \right) \right) + \left(\left(-\frac{1}{8} \right) \cdot \log_2 \left(\frac{1}{8} \right) \right) \\ &= 1,4056 \end{aligned}$$

Secara lengkap hasil perhitungan nilai *entropy* masing-masing atribut ke- i disajikan pada Tabel 4.

Tabel 4 Nilai *Entropy* Masing-Masing Atribut Ke- i

Atribut	<i>Entropy</i>	Atribut	<i>Entropy</i>	Atribut	<i>Entropy</i>
A_1	0,971	E_1	1,4655	I_2	0,7219
A_2	1,4056	E_2	1	J_1	0,7219
A_3	1,4837	F_1	1,51	J_2	0
B_1	1,4655	F_2	0,9911	J_3	1,2997
B_2	1	G_1	1,361	J_4	1,5
C_1	1,4605	G_2	1,4855	K_1	1,4406
C_2	0,8113	H_1	1,4197	K_2	1,2955
D_1	1,4488	H_2	0,8631	L_1	1,5
D_2	1,4282	I_1	1,4619	L_2	1,314

Berdasarkan informasi Tabel 4, setelah mendapat nilai *entropy* masing-masing atribut ke-*i* kemudian dicari nilai *gain* masing-masing atribut. Perhitungan nilai *gain* masing-masing atribut dengan *entropy* total = 1,4332. Dimulai dengan atribut *A* dengan $\sum|A|$ adalah jumlah responden balita masing-masing kelompok atribut *A* dan *N* adalah total seluruh responden balita. *Entropy A* adalah kumpulan masing-masing nilai *entropy* pada atribut *A*.

Diperoleh nilai *gain* :

$$\begin{aligned} Gain(A) &= Entropy\ Total - \left(\sum_{j=1}^k \frac{|A_j|}{|N|} \cdot Entropy\ A_j \right) \\ &= 1,4332 - \left(\left(\frac{5}{30} \cdot 0,971 \right) + \left(\frac{8}{30} \cdot 1,4056 \right) + \left(\frac{17}{30} \cdot 1,4837 \right) \right) \\ &= 0,0558 \end{aligned}$$

Dengan *entropy* total = 1,4332, $\sum|B|$ adalah jumlah responden balita masing-masing kelompok atribut *B* dan *N* adalah total seluruh responden balita. *Entropy B* adalah kumpulan masing-masing nilai *entropy* pada atribut *B*.

Diperoleh nilai *gain* :

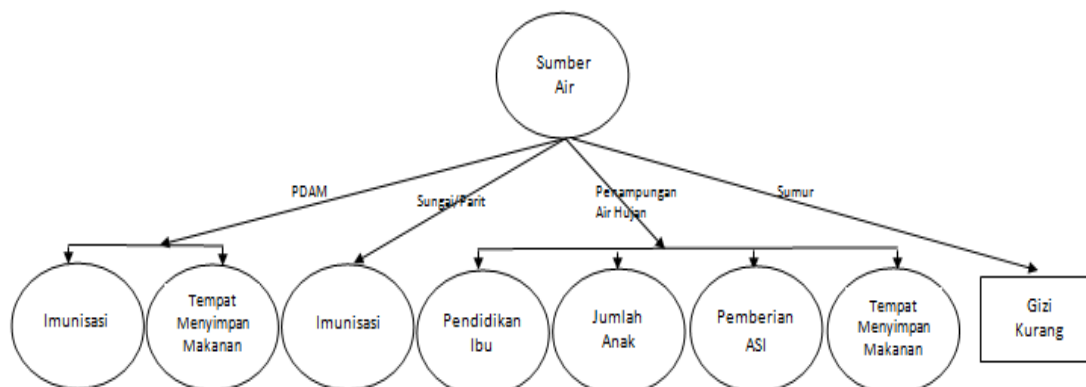
$$\begin{aligned} Gain(B) &= Entropy\ Total - \left(\sum_{j=1}^k \frac{|B_j|}{|N|} \cdot Entropy\ B_j \right) \\ &= 1,4332 - \left(\left(\frac{26}{30} \cdot 1,4655 \right) + \left(\frac{4}{30} \cdot 1 \right) \right) \\ &= 0,0297 \end{aligned}$$

Secara lengkap hasil perhitungan nilai *gain* masing-masing atribut disajikan pada Tabel 5.

Tabel 5 Perhitungan Nilai *Gain* Masing-Masing Atribut

Atribut	<i>Gain</i>	Atribut	<i>Gain</i>
G_A	0,0558	G_H	0,2731
G_B	0,0297	G_I	0,0946
G_C	0,0593	G_J	0,3330
G_D	0,0002	G_K	0,0409
G_E	0,0297	G_L	0,0940
G_F	0,0788	Gain Terbesar G_J	0,3330
G_G	0,0307		

Berdasarkan informasi Tabel 5, atribut yang memiliki nilai *gain* terbesar yakni atribut yaitu atribut J. Atribut J merupakan atribut sumber air yang digunakan. Oleh karena itu, atribut J dijadikan sebagai simpul akar. Atribut J ini memiliki 4 atribut, yaitu penampungan air hujan, sumur, sungai dan PDAM. Keempat atribut ini digunakan sebagai cabang. Dari keempat atribut ini dilihat kembali nilai *entropy*-nya. Atribut yang memiliki nilai *entropy* terbesar akan berada disebelah kiri dan seterusnya. Perhitungan dan pembentukan simpul akan berhenti apabila atribut telah menghasilkan sebuah solusi dari kasus ketiga yang diteliti. Secara umum, pohon keputusan tingkat I dapat dilihat pada Gambar 1.



Gambar 1 Pohon Keputusan

Gambar 1 memperlihatkan bahwa atribut sumber air yang digunakan menjadi akar dari pohon keputusan karena memiliki nilai *gain* terbesar. Atribut PDAM berada paling kiri karena memiliki nilai *entropy* terbesar yaitu 1,5, kemudian atribut sungai/parit dengan nilai *entropy* terbesar kedua yaitu 1,2997. Selanjutnya, atribut penampungan air hujan dengan nilai *entropy* sebesar 0,7219 sedangkan atribut sumur memiliki nilai *entropy* terkecil yaitu 0. Setelah membentuk simpul akar, maka pembentukan pohon dilanjutkan dengan menghitung nilai *entropy* dan *gain* pada masing-masing atribut. Perhitungan berakhir hingga menemukan atribut dengan nilai *entropy* 0. Dengan kata lain simpul tersebut tidak dapat membentuk cabang dibawahnya.

PENUTUP

Algoritma C4.5 merupakan salah satu metode dalam pohon keputusan yang mengklasifikasikan variabel-variabel yang memiliki keterkaitan dengan melihat nilai *gain* tertinggi. Penerapan algoritma dalam melihat status gizi balita Desa Kuala Dua Kabupaten Kubu Raya menghasilkan pohon keputusan dengan 2 tingkatan. Akar pohon terbentuk oleh variabel sumber air bersih yang digunakan dengan nilai *gain* terbesar yaitu 0,3330. Variabel sumber air bersih yang digunakan terdiri dari atribut penampungan air hujan, sungai/parit, sumur dan PDAM.

Variabel yang memiliki berpengaruh selanjutnya dalam status gizi balita yaitu tingkat I pohon terdiri dari variabel pelaksanaan imunisasi dengan nilai *gain* sebesar 1, variabel tempat menyimpan makanan dengan nilai *gain* sebesar 0,2763, variabel pelaksanaan imunisasi dengan nilai *gain* sebesar 0,3219. Kemudian variabel-variabel yang memiliki nilai *gain* yang sama yaitu sebesar 0,3219 diantaranya variabel pendidikan ibu, jumlah anak yang dimiliki, variabel pemberian ASI, tempat menyimpan makanan.

Pembentukan pohon tingkat II dibentuk oleh variabel dengan dengan nilai *gain* sebesar 1 yaitu variabel pemberian ASI sejak lahir, variabel intensitas makan balita, variabel porsi makan balita dan variabel pengolahan makanan. Selanjutnya variabel-variabel yang memiliki nilai *gain* yang sama yaitu sebesar 1,557 diantaranya variabel pemberian ASI sejak lahir, variabel intensitas makan balita, variabel porsi makan balita dan variabel pengolahan makanan. Adapun variabel-variabel selanjutnya yaitu variabel pendidikan ibu, variabel jumlah anak yang dimiliki, variabel pemberian ASI, dan variabel tempat menyimpan makanan dengan nilai *gain* sebesar 0,3219.

Oleh karena variabel-variabel yang terdapat pada tingkat II telah menghasilkan solusi akan kasus yang diteliti, maka pembentukan pohon berakhir pada tingkat II.

DAFTAR PUSTAKA

- [1]. Irianto, K. *Gizi Seimbang dalam Kesehatan Reproduksi*. Bandung : Alfabeta; 2014.
- [2]. Adisasmito, W. *Sistem Kesehatan*. Jakarta : PT. Rajagrafindo Persada; 2012.
- [3]. Badan Pusat Statistik. *Kabupaten Kubu Raya Dalam Angka*. Kabupaten Kubu Raya : BPS; 2014.
- [4]. Puskesmas Sungai Durian. *Profil Puskesmas Sungai Durian*. Kabupaten Kubu Raya : Puskesmas Sungai Durian; 2015.
- [5]. Kusriani, dan Luthfi, E. T. *Algoritma Data Mining*. Yogyakarta : Andi; 2009.
- [6]. Kemenkes RI No. 1995/MENKES/SK/XII/2010 *Tentang Standar Antropometri Penilaian Status Gizi Anak*. Jakarta : Menkes; 2011.
- [7]. Munir, R. *Matematika Diskrit*. Edisi Ketiga. Bandung : Informatika; 2005
- [8]. Zakaria, Marcus, T, Agus, P. *Konsep dan Implementasi Struktur Data*. Bandung : Informatika; 2006

SUPRIYATI

: FMIPA Untan Pontianak, supriyati_mathematic09@ymail.com

HENDRA PERDANA

: FMIPA Untan Pontianak, hendra.perdana@math.untan.ac.id