

PENENTUAN TINGKAT KESEJAHTERAAN ANAK MENGUNAKAN ALGORITMA C 4.5

Yuli Murdianingsih¹, Abdul Syukur², M Arief Soeleman³
^{1,2,3}Pasca Sarjana Teknik Informatika Universitas Dian Nuswantoro

ABSTRACT

Realization of child welfare is a right of every child and is the responsibility of all. Today the program services are sporadic, discontinue and responsiveness is a result of not optimal data management with social welfare problems is very large. Need a model system that can help make decisions quickly, precisely and accurately. In this research the basic needs of children based on four parameters: physical, intellectual, emotional, social and spiritual. C4.5 algorithm implemented in the m system's model of children's basic needs level is done by calculating the entropy and the gain of the parameters of physical, intellectual, emotional and spiritual social iteratively in order to obtain a decision tree and rules used to model. Data analysis base on 149 datas as the training data and the testing data is 37. The accuracy of the model to look at the performance of the system using confusion matrix. Systems decision trees obtained the degree of basic needs of children, from the decision tree obtained seven rules that are used in view, values of accuracy obtained 94,59 % . C 4.5 algorithm can be used for classification of the level of a child's basic needs are met and not met.

Keywords: Basic needs of children, C4.5, entropy, gain, accuracy

1. PENDAHULUAN

1.1. Latar Belakang

Dalam era globalisasi, diperlukan sumber daya manusia yang berkualitas agar mampu bersaing. Untuk mendapatkannya harus dimulai sejak dini, tidak hanya dalam pertumbuhan fisik saja tetapi juga dalam perkembangan mental, sosial dan emosional. Sehingga kelak akan menjadi manusia yang mempunyai kecerdasan intelektual, kecerdasan emosional, kecerdasan sosial dan kecerdasan mental-spiritual yang tinggi [1]. Untuk mendapatkan anak seperti yang diharapkan tersebut di atas, tentunya diperlukan berbagai upaya baik oleh keluarga, masyarakat maupun pemerintah. Bila sejak awal sudah didapatkan bibit-bibit yang baik dan diberikan lingkungan yang baik, diharapkan tumbuh kembang anak akan optimal seperti yang diharapkan sehingga perlindungan dan pemenuhan tingkat kesejahteraan anak secara umum harus mendapat perhatian yang memadai. Usaha kesejahteraan anak terdiri atas usaha pembinaan, pengembangan, pencegahan, dan rehabilitasi [2].

Berpijak pada hal tersebut, maka pemenuhan kebutuhan anak sebagai salah satu indikator pencapaian kesejahteraan anak harus diupayakan dan terus ditingkatkan baik dalam bentuk pelayanan maupun penyuluhan. Usaha untuk mewujudkan kesejahteraan anak mengalami kendala baik teknis maupun struktural. Kendala ini terutama dialami oleh sebagian wilayah yang letaknya jauh dari pusat. Kota Bandung khususnya kecamatan cicadas merupakan daerah dengan kondisi terkumuh di kota Bandung, dimana kekumuhan merupakan salah satu indikator kemiskinan. Sebagian besar permasalahan kesejahteraan social disebabkan oleh kemiskinan [3], hal serupa juga disampaikan oleh Kadushin [4] bahwa ketidakmampuan orang tua dalam usaha kesejahteraan anak adalah kemiskinan yang menyelubungi mereka. Departemen Sosial menangani 19 jenis penyandang masalah kesejahteraan sosial (PMKS) yang diantaranya berkaitan dengan masalah pemenuhan kebutuhan dasar dan hak anak antara lain : anak balita terlantar, anak terlantar, anak nakal, anak jalanan, penyandang cacat dan penyalahgunaan napza [3].

Dengan menggunakan parameter tingkat pemenuhan kebutuhan anak berupa fisik, emosional, intelektual dan sosial spiritual diharapkan dapat mengungkap deskripsi tingkat pencapaian kesejahteraan anak.

1.2. Rumusan Masalah

- a. Belum ada pengelolaan data tingkat kesejahteraan anak khususnya pemenuhan kebutuhan dasar anak sehingga pelayanan bersifat sporadis, diskontinu, dan responsif.
- b. Pemetaan penyebab retardasi mental pada anak secara klinis dengan algoritma C4.5 akurasi masih kurang optimal.

1.3. Tujuan

- a. Memberikan masukan untuk pengelolaan data tingkat kesejahteraan anak khususnya pemenuhan kebutuhan dasar anak agar pelayanan tidak bersifat sporadis, diskontinu, dan responsif.
- b. Diperolehnya pemetaan penyebab retardasi mental pada anak secara klinis dengan akurasi yang lebih tinggi daripada penelitian sebelumnya.

2. LANDASAN TEORI

2.1. Algoritma C 4.5

Craw [6] menyampaikan pohon keputusan menggunakan algoritma C 4.5 dalam menentukan keputusan bermain tenis berdasarkan pada outlook, temperature, humidity dan windy. Untuk memperoleh root pada pohon keputusan dan seluruh node yang lain yang tertera pada pohon keputusan, Craw melakukan penghitungan *Entropy*, seperti pada Rumus 1, melakukan penghitungan Gain informasi seperti pada Rumus 2.

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

Keterangan:
 S : himpunan kasus
 n : jumlah partisi S
 pi : proporsi dari Si terhadap S

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i) \dots\dots\dots 2)$$

Keterangan:
 S : himpunan kasus
 A : atribut
 n : jumlah partisi atribut A
 |Si| : jumlah kasus pada partisi ke-i
 |S| : jumlah kasus dalam S

Proses klasifikasi meliputi dua tahap, yaitu tahap pertama adalah proses *learning* yaitu data *training* dianalisis oleh algoritma klasifikasi dan menghasilkan sejumlah aturan, yaitu proses menganalisis data *training* oleh algoritma klasifikasi, kemudian tahap kedua yaitu *training* adalah menguji akurasi dari algoritma klasifikasi dengan memasukkan sejumlah data [7].

- a. Menyiapkan data *training*. Data *training* diambil dari penelitian dan sudah dikelompokkan ke dalam kelas-kelas tertentu.
- b. Menentukan akar dari pohon. Akar akan diambil dari atribut yang terpilih, dengan cara menghitung nilai Gain dari masing-masing atribut, nilai Gain yang paling tinggi yang akan menjadi akar pertama. Sebelum menghitung nilai Gain dari atribut, terlebih dahulu dihitung nilai *entropy*.
- c. Kemudian hitung nilai Gain dengan metode *information gain*.

- d. Ulangi langkah ke-2 hingga semua atribut terpartisi.
- e. Proses partisi pohon keputusan akan berhenti saat :
 - 1) Semua atribut dalam node n mendapat kelas yang sama.
 - 2) Tidak ada atribut di dalam *tuple* yang dipartisi lagi.
 - 3) Tidak ada atribut di dalam cabang yang kosong.

Demikian langkah-langkah dalam implementasi algoritma C 4.5, sehingga diperoleh pohon keputusan yang lengkap baru dilakukan ekstraksi pada aturan yang harus ada dalam penentuan tingkat pemenuhan kebutuhan dasar anak.

2.2. Klasifikasi Rule based

Rule based atau algoritma berbasis aturan merupakan cara terbaik untuk merepresentasikan sejumlah bit data atau pengetahuan [7]. *Rule based* biasanya dituliskan dalam bentuk logika *IF-THEN* atau jika dibuat persamaannya yaitu :

IF condition THEN conclusion

Pernyataan IF dari persamaan di atas dikenal sebagai *rule antecedent* atau precondition sedangkan pernyataan THEN disebut sebagai *rule consequent*. Dalam *rule antecedent* biasanya menyertakan satu atau lebih atribut dan menggunakan logika AND jika menggunakan lebih dari satu atribut. *Rule consequent* merupakan prediksi kelas [7].

Aturan-aturan dalam *Rule based* dapat diturunkan dari pohon keputusan yang telah terbentuk. Karena pohon keputusan yang besar, terkadang sulit untuk menginterpretasikan pohon bentuk keputusan [7]. Agar pohon keputusan ini dapat lebih mudah dipahami oleh manusia, maka perlu diinterpretasikan dalam bentuk aturan-aturan atau *Rule based*.

2.3. Confusion Matrix

Confusion matrix [8] merupakan metode untuk mengevaluasi model klasifikasi pada *data mining* dengan menghasilkan nilai prediksi benar dan prediksi salah jika dibandingkan ke nilai tujuan (*target value*) dalam data. *Confusion matrix* adalah suatu metode yang biasanya digunakan untuk melakukan perhitungan akurasi pada konsep *data mining*. Rumus ini melakukan perhitungan dengan lima keluaran, yaitu: *recall*, *specificity*, *precision*, *negative predictive value* dan *accuracy*.

Tabel 1. *Confusion Matrix*

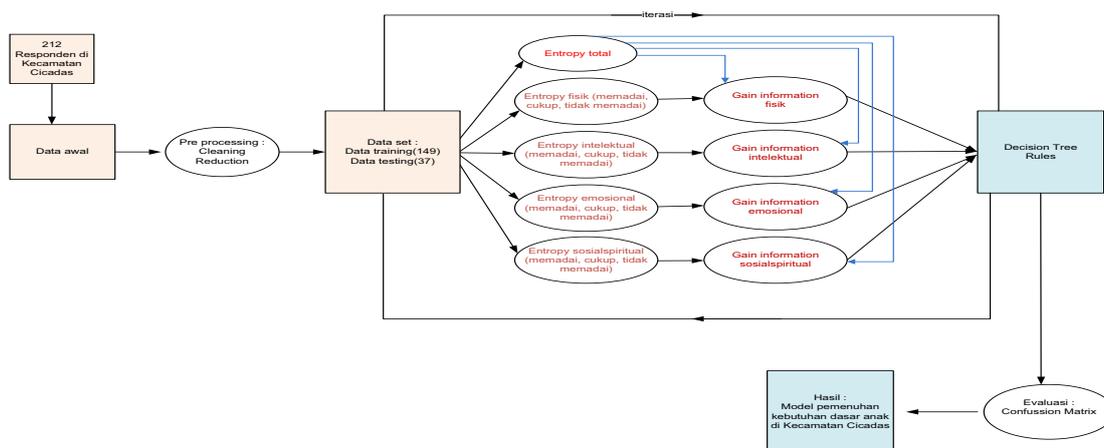
Confusion Matrix		Target			
		Positive	Negative		
Model	Positive	a	b	<i>Positive Predictive Value</i>	$a/(a+b)$
	Negative	c	d	<i>Negative Predictive Value</i>	$d/(c+d)$
		<i>Sensitivity</i>	<i>Specificity</i>	Accuracy = $(a+d)/(a+b+c+d)$	
		$a/(a+c)$	$d/(b+d)$		

3. PENENTUAN TINGKAT PEMENUHAN KESEJAHTERAAN ANAK BERDASARKAN ALGORITMA C4.5

Berdasarkan pada Hurlock [1] tingkat pemenuhan kebutuhan anak meliputi parameter kebutuhan fisik, intelektual, emosional dan sosial spiritual sedangkan indikator pemenuhan kesejahteraan anak mengacu pada indikator yang telah ditentukan oleh BKKBN [9] dan dijadikan acuan dalam pembuatan instrumen

penelitian yang terdiri dari: **Kebutuhan Fisik, Kebutuhan Intelektual, Kebutuhan Sosial Spiritual, Kebutuhan Emosional.**

Penentuan tingkat kesejahteraan anak yang menggunakan algoritma C4.5 dapat diartikan sebuah upaya menentukan tingkat kesejahteraan anak melalui pohon keputusan yang diperoleh dengan cara melakukan penghitungan *entropy* total, *entropy* fisik untuk nilai memadai, cukup, tidak memadai, *entropy* intelektual untuk nilai memadai, cukup, tidak memadai, *entropy* emosional untuk nilai memadai, cukup, tidak memadai dan *entropy social spiritual* untuk nilai memadai, cukup, tidak memadai kemudian menentukan gain informasi terhadap parameter-parameter sumber berupa parameter kebutuhan fisik, intelektual, emosional dan sosial spiritual terhadap parameter tujuan atau class dengan dua kategori *class* yaitu terpenuhi dan tidak terpenuhi.



Gambar 1. Kerangka Pikir Penelitian

4. METODE PENELITIAN

4.1. Teknik Pengambilan Data

Pengambilan sampel pada penelitian ini menggunakan metode *Non-Probability Sampling* dengan teknik *incidental sampling*. Dengan demikian, tidak semua elemen dalam populasi memiliki kesempatan yang sama untuk dipilih sebagai sampel penelitian. Subyek dipilih berdasarkan ketersediaan dan kesediaan mereka. Selain alasan kemudahan, teknik ini juga sederhana pelaksanaannya. Teknik pengumpulan data dalam penelitian ini menggunakan kuesioner, yaitu suatu daftar berisikan pertanyaan-pertanyaan tentang suatu hal atau bidang tertentu yang telah disusun untuk memperoleh data berupa jawaban tertulis dari partisipan yang dijadikan sampel [10].

4.2. Pengolahan Data

a. *Cleaning & Integration*

Tahap ini merupakan proses membersihkan nilai yang kosong atau data yang tidak lengkap atau tidak sesuai dengan kriteria penelitian. Setelah melalui proses *cleaning*, dari 212 data responden yang dikumpulkan diperoleh 186 responden yang sesuai dengan kriteria dan digunakan sebagai sampel, 80% atau sebanyak 149 dijadikan data *training* dan 20% atau 37 sebagai data *testing*.

b. *Reduction*

Jumlah atribut yang digunakan untuk *data mining* mungkin terlalu besar sehingga beberapa atribut yang tidak digunakan akan dihapus. Setelah melalui proses *reduction*, dalam penelitian ini terdapat

beberapa atribut dalam karakteristik responden yang berisi no responden, nama, usia, alamat, pendidikan, pekerjaan, pendapatan, status pernikahan, jumlah anak dan jumlah anggota keluarga yang sudah bekerja. Jumlah anggota keluarga yang sudah bekerja tidak digunakan karena pendapatan dianggap telah mewakili atribut tersebut. Tabel 2 menunjukkan sampel data set setelah dilakukan pre processing sebanyak 186 responden.

Tabel 2. Data Set setelah *Preprocessing*

NO	NAMA	FISIK	EMOSIONAL	SOSIALSPIRITUAL	INTELEKTUAL	JUMLAH
1	AJI	TIDAK MEMADAI	CUKUP	MEMADAI	CUKUP	TIDAK TERPENUHI
2	MULYANA	CUKUP	MEMADAI	MEMADAI	CUKUP	TIDAK TERPENUHI
3	ASEP	CUKUP	CUKUP	MEMADAI	CUKUP	TERPENUHI
4	MEGAWATI	TIDAK MEMADAI	MEMADAI	MEMADAI	MEMADAI	TERPENUHI
5	AANG	CUKUP	MEMADAI	MEMADAI	MEMADAI	TERPENUHI
6	ADE	MEMADAI	MEMADAI	MEMADAI	CUKUP	TERPENUHI
7	ASTRI	CUKUP	MEMADAI	MEMADAI	CUKUP	TERPENUHI
8	SITI	CUKUP	MEMADAI	MEMADAI	MEMADAI	TERPENUHI
9	DODO	CUKUP	MEMADAI	MEMADAI	CUKUP	TERPENUHI
10	IPAN	MEMADAI	MEMADAI	MEMADAI	CUKUP	TERPENUHI
11	HERI	MEMADAI	MEMADAI	MEMADAI	CUKUP	TERPENUHI
12	DARLIAH	TIDAK MEMADAI	MEMADAI	MEMADAI	MEMADAI	TERPENUHI
13	IRMA	CUKUP	MEMADAI	MEMADAI	MEMADAI	TERPENUHI
14	LALAN SUHERLAN	CUKUP	MEMADAI	MEMADAI	CUKUP	TERPENUHI
15	APUD	TIDAK MEMADAI	MEMADAI	CUKUP	CUKUP	TIDAK TERPENUHI
16	HANAN	MEMADAI	MEMADAI	MEMADAI	MEMADAI	TERPENUHI
17	KUSNADI	CUKUP	MEMADAI	MEMADAI	CUKUP	TIDAK TERPENUHI
18	TINI	TIDAK MEMADAI	MEMADAI	MEMADAI	MEMADAI	TIDAK TERPENUHI
19	SOMANTRI	CUKUP	MEMADAI	MEMADAI	CUKUP	TERPENUHI
20	ENO	CUKUP	MEMADAI	MEMADAI	MEMADAI	TERPENUHI

4.3. Metode *Scoring*

Jumlah pertanyaan dalam kuesioner penelitian cukup banyak sehingga diperlukan *scoring* untuk memudahkan dalam proses penilaian dan akan membantu dalam proses analisis data yang telah ditemukan. Penentuan *scoring* ilmiah secara umum berpedoman pada aturan Likert dan Gutman [11]. Pada penelitian ini, kuesioner diukur menggunakan skala likert dengan tiga pilihan jawaban dengan skor 1 untuk nilai jawaban terendah dan skor 3 untuk nilai jawaban tertinggi. Penentuan *scoring* pada penilaian objektif dilakukan untuk menentukan rentang atau nilai interval dalam mengubah data kedalam bentuk kategorikal. Berdasarkan kuesioner penilaian dan pemberian *scoring* dengan menggunakan pendekatan skala likert terlihat pada Tabel 3 berikut.

Tabel 3. Perhitungan Skala Likert

JENIS KEBTHAN	PERHITUNGAN SKALA LIKERT	KATEGORI
FISIK	Skor tertinggi = 23 = 100%	MEMADAI ≥ 18 CUKUP = 13-17 TIDAK MEMADAI ≤ 12
	Skor terendah = $8/23 \times 100 = 34.78\%$	
	Range = $100 - 34.78 = 65.22\%$	
	Kategori = 3	
	Interval = $65.22/3 = 21.74\% \times 23 = 5$	
EMOSIONAL	Skor tertinggi = 12 = 100%	MEMADAI ≥ 9 CUKUP = 6-8 TIDAK MEMADAI ≤ 5
	Skor terendah = $4/12 \times 100 = 33,33\%$	
	Range = $100 - 33,33 = 66,67$	
	Kategori = 3	
	Interval = $66,67/3 = 22,22\% \times 12 = 2,67$	
INTELEKTUAL	skor tertinggi = 9 = 100%	MEMADAI ≥ 7 CUKUP = 5-6 TIDAK MEMADAI ≤ 4
	Skor terendah = $3/9 \times 100\% = 33.33\%$	
	Range = $100 - 33.33\% = 66.67$	
	Kategori = 3	
	Interval = $66.67/3 = 22.22\% \times 9 = 1.99$	
SOSIAL SPIRITUAL	Skor tertinggi = 12 = 100%	MEMADAI ≥ 9 CUKUP = 6-8 TIDAK MEMADAI ≤ 5
	Skor terendah = $4/12 \times 100\% = 33,33\%$	
	Range = $100 - 33,33 = 66,67$	
	Kategori = 3	
	Interval = $66,67/3 = 22,22\% \times 12 = 2,67$	
CLASS	Skor tertinggi = 54 = 100%	TERPENUHI $\geq 36,50$ TIDAK TERPENUHI $< 36,50$
	Skor terendah = $19/54 = 0,352 \times 100\% = 35.2\%$	
	Range = $100 - 35.2 = 64.8\%$	
	Kategori = 2 (terpenuhi, tidak terpenuhi)	
	Interval = $64.8/2 = 32.4\%$	
Criteria penilaian = $100 - 32.4 = 67.6\% \times 54 = 36,50$		

4.4. Penentuan Data Training dan Data Testing

Dari 186 data responden, dibagi menjadi 80% untuk *training* dan 20% sisanya sebagai *testing*. Jumlah data untuk *training* dan *testing* [12] masing-masing adalah :

$N_{training} = 80/100 \times 186 = 149$ responden

$N_{testing} = 20/100 \times 186 = 37$ responden

Sedangkan proporsi masing-masing kategori (terpenuhi dan tidak terpenuhi) dalam data keseluruhan, adalah:

$$P_1 = \frac{159}{186} \times 100 = 85,48\%$$

$$P_2 = \frac{27}{186} \times 100 = 14,52\%$$

Proporsi masing-masing kategori dalam data *training* dan data *testing* harus sesuai dengan proporsi masing-masing kategori dalam data keseluruhan. Sehingga jumlah masing-masing kategori dalam data *training* dan data *testing* adalah :

Untuk data *training*

$$N_1 = P_1 \times N_{\text{training}} = \frac{85,48}{100} \times 149 = 127,37 \approx 127$$

$$N_2 = P_2 \times N_{\text{training}} = \frac{14,52}{100} \times 149 = 21,63 \approx 22$$

Untuk data *testing*

$$N_1 = P_1 \times N_{\text{testing}} = \frac{85,48}{100} \times 37 = 31,63 \approx 32$$

$$N_2 = P_2 \times N_{\text{testing}} = \frac{14,52}{100} \times 37 = 5,37 \approx 5$$

Kategori 1 untuk Class Terpenuhi = $127 + 32/32 = 159/32 = 4,96 \approx 5$

Kategori 2 untuk Class Tidak terpenuhi = $22 + 5/5 = 27/5 = 5,4 \approx 5$

Dengan demikian untuk data kategori 1 (terpenuhi) sebanyak 159 data, diambil satu secara acak dari lima data sebagai data *testing* sebanyak 32 data. Untuk data kategori 2 (tidak terpenuhi), diambil sebanyak 5 secara acak dari 27 data yang ada sebagai data *testing*, sehingga data *testing* keseluruhan berjumlah 37 responden.

5. HASIL PENELITIAN DAN PEMBAHASAN

$$\begin{aligned} Entropy(S) &= \sum_{i=1}^n -p_i * \log_2 p_i \\ &= -p_1 * \log_2 p_1 - p_2 * \log_2 p_2 \\ &= -\left(\frac{S_1}{S}\right) * \log_2 \left(\frac{S_1}{S}\right) - \left(\frac{S_2}{S}\right) * \log_2 \left(\frac{S_2}{S}\right) \\ &= -\frac{128}{149} * \log_2 \left(\frac{128}{149}\right) - \frac{22}{149} * \log_2 \left(\frac{22}{149}\right) \\ &= 0.188279 + 0.407478 \\ &= 0.595757 \end{aligned}$$

Entropy untuk atribut fisik ($S_{memadai}$, S_{cukup} , $S_{tidak\ memadai}$):

$$\begin{aligned} Entropy(S_{memadai}) &= \sum_{i=1}^n -p_i * \log_2 p_i \\ &= -\frac{96}{103} * \log_2 \left(\frac{96}{103}\right) - \frac{7}{103} * \log_2 \left(\frac{7}{103}\right) \\ &= 0.094637 + 0.263631 \\ &= 0.358269 \end{aligned}$$

$$\begin{aligned} Entropy(S_{cukup}) &= \sum_{i=1}^n -p_i * \log_2 p_i \\ &= -\frac{10}{15} * \log_2 \left(\frac{10}{15}\right) - \frac{5}{15} * \log_2 \left(\frac{5}{15}\right) \\ &= 0.389975 + 0.528321 \\ &= 0.918296 \end{aligned}$$

$$\begin{aligned} Entropy(S_{tidakmemadai}) &= \sum_{i=1}^n -p_i * \log_2 p_i \\ &= -\frac{0}{32} * \log_2 \left(\frac{0}{32}\right) - \frac{32}{32} * \log_2 \left(\frac{32}{32}\right) \\ &= 0 + 0 \\ &= 0 \end{aligned}$$

Information Gain untuk nilai fisik:

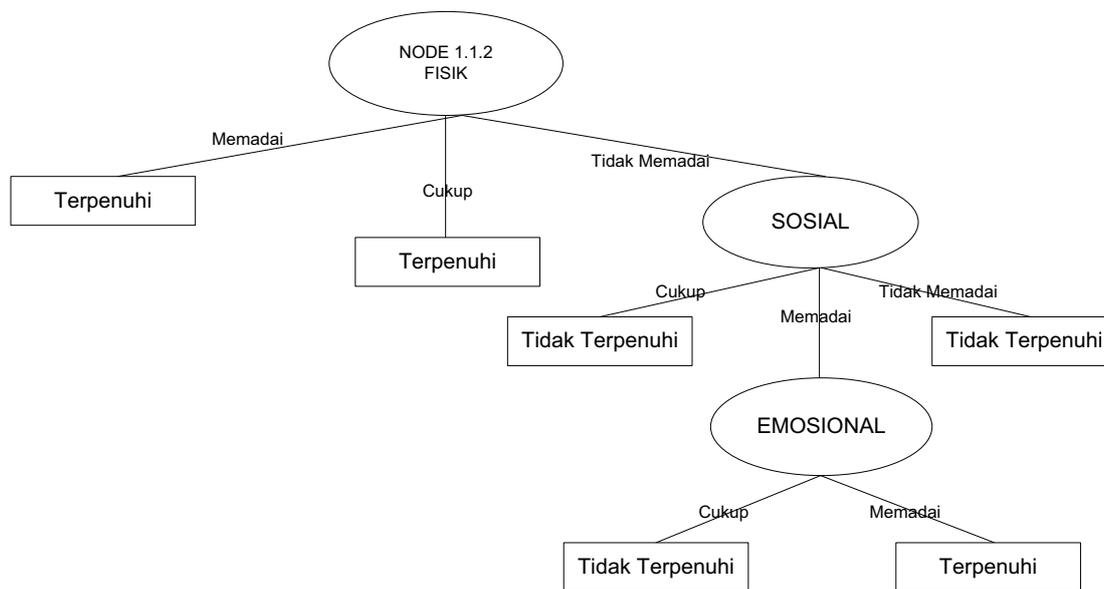
$$\begin{aligned} Gain(S, A) &= Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i) \\ &= Entropy(total) - \sum_{i=1}^3 \frac{S_{fisik}}{S_{total}} * Entropy(fisik) \\ &= 0.872988 - \left(\frac{103}{150}\right) * 0.358269 - \left(\frac{15}{150}\right) * 0.918296 - \left(\frac{32}{150}\right) * 0 \\ &= 0.535148 \end{aligned}$$

Demikian juga untuk atribut yang lainnya dilakukan proses perhitungan serupa, sehingga pada fase pencarian akar diperoleh rangkuman hasil perhitungan seperti diperlihatkan dalam Tabel 4.

Tabel 4. Rangkuman Hasil Perhitungan *Entropy* dan *Gain* untuk *Root* Pohon Keputusan

Node	Atribut	Jml Kasus S	Terpenuhi	tidak Terpenuhi	<i>Entropy</i>	<i>Gain</i>
1	Total	149	128	2	0.595757	
	fisik					0.286768
	Memadai	89	89		0	
	Cukup	27	24		0.503258	
	tidak memadai	33	14	9	0.983376	
	intelektual					0.073316
	Memadai	78	75		0.235193	
	Cukup	71	52	9	0.838008	
	tidak memadai	0	0		0	
	emosional					0.051488
	Memadai	142	125	7	0.528551	
	Cukup	7	2		0.863121	
	tidak memadai	0	0		0	
	sosial spiritual					0.016847
	Memadai	132	116	6	0.532835	
	Cukup	17	11		0.936667	
	tidak memadai	0	0		0	

Setelah semua atribut dalam node n mendapat kelas yang sama, tidak ada atribut di dalam *tuple* yang dipartisi lagi dan tidak ada atribut di dalam cabang yang kosong, maka perhitungan *entropy* dan *gain* selesai dan diperoleh pohon keputusan seperti pada Gambar 2.



Gambar 2. Pohon Keputusan

Berdasarkan pohon keputusan di atas, diperoleh 6 aturan sebagai berikut :

R1: IF fisik=memadai THEN class=terpenuhi

R2: IF fisik=cukup THEN class=terpenuhi

R3: IF fisik=tidak memadai AND sosial spiritual=cukup THEN class= tidak terpenuhi

R4: IF fisik=tidak memadai AND sosial spiritual=tidak memadai THEN class= tidak terpenuhi

R5: IF fisik=tidak memadai AND sosial spiritual=memadai AND emosional=cukup THEN class= tidak terpenuhi

R6: IF fisik=tidak memadai AND sosial spiritual=memadai AND emosional=memadai THEN class= terpenuhi

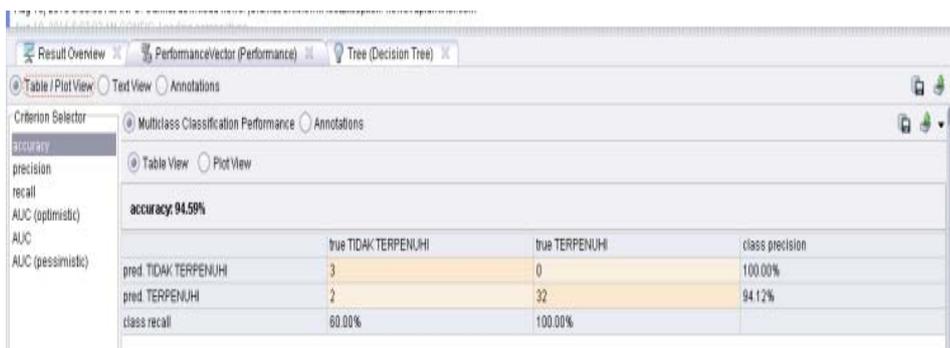
Hasil Pengujian dan Evaluasi

Untuk menguji akurasi dari model yang telah terbentuk, digunakan metode *Confusion Matrix*. Metode ini biasanya digunakan untuk melakukan perhitungan akurasi pada konsep *data mining*. Langkah pertama adalah menghitung kecocokan pada kelas prediksi dengan target atau kelas actual dengan membandingkan data *testing* pada pohon keputusan, sehingga diperoleh lima keluaran *Confusion Matrix* sebagai berikut :

Tabel 5. Perhitungan *Confusion Matrix*

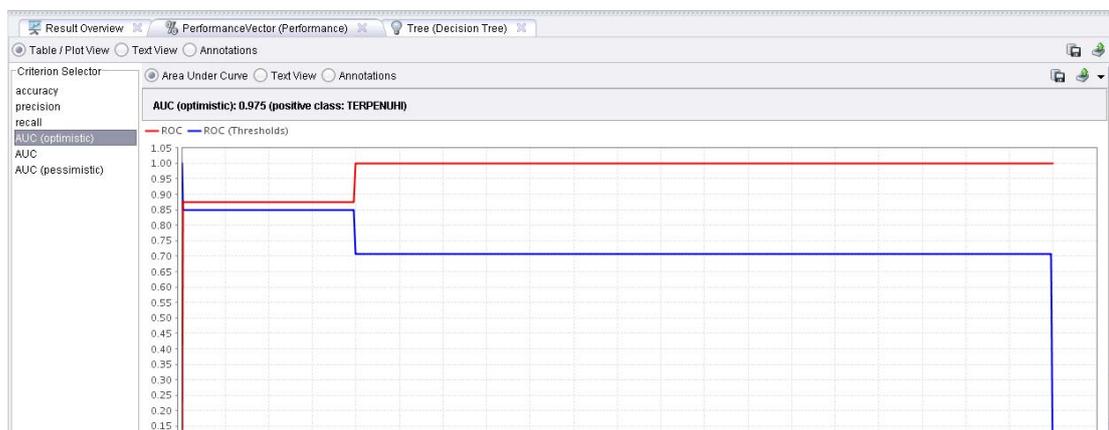
<i>Confusion Matrix</i>		Target			
		positive	negative		
Model	Positive	32	0	Positive predictive value	32/32=100%
	negative	2	3	Negative predictive value	3/5=60%
		Sensitivity	Specificity	Accuracy = 35/37=94,59%	
		32/34=94,12%	3/3=100%		

Hasil perhitungan *confussion matrix* menunjukkan bahwa nilai akurasi yang diperoleh dalam upaya pemenuhan kebutuhan dasar anak di Kecamatan Cicadas menggunakan algoritma C4.5 adalah 94,59%, hal ini menunjukkan bahwa model yang diperoleh memiliki nilai kehandalan yang memadai. Gambar 3 dan 4. nilai *accuracy*, *recall* dan *precision* yang dihasilkan oleh Rapid Miner menggunakan model *confussion matrix*, berdasarkan gambar tersebut, kita bisa lihat adanya kesesuaian antara nilai keluaran *confussion matrix* hasil perhitungan manual dengan nilai yang dihasilkan oleh software *Rapid Miner*.



Gambar 3. Hasil Akurasi Berdasarkan Rapid Miner

Kurva ROC (Receiver Operating Characteristic)



Gambar 4. Nilai AUC dalam Grafik ROC dengan Metode Information Gain

Gambar 4 menunjukkan grafik ROC dengan nilai AUC (*Area Under Curve*) dengan metode *information gain* sebesar 0.975. Akurasi AUC dikatakan sempurna apabila nilai AUC mencapai 1.000 dan akurasinya buruk jika nilai AUC di bawah 0.500. dengan demikian nilai AUC yang dihasilkan termasuk dalam tingkat diagnos *Good Classification*.

6. KESIMPULAN DAN SARAN

6.1 Kesimpulan

- a. Dari hasil penelitian ini diperoleh model tingkat pemenuhan kebutuhan dasar anak menggunakan algoritma C4.5 dan dapat dilihat bahwa pola pemenuhan kebutuhan dasar anak di Kecamatan

Cicadas adalah pemenuhan kebutuhan fisik, pemenuhan kebutuhan social spiritual dan pemenuhan kebutuhan emosional.

- b. Akurasi yang diperoleh dengan menggunakan model *confusion matrix* lebih besar dari akurasi yang diperoleh pada penelitian sebelumnya, yaitu sebesar 94,59% dan nilai ini dianggap memiliki kehandalan yang memadai dan dari grafik ROC dihasilkan nilai AUC sebesar 0,975 dengan tingkat diagnose *Good Classification*.
- c. Kemungkinan *rule* yang diperoleh dari data penelitian ini adalah 4(atribut) x 12 (instance) sehingga kemungkinan aturan yang diperoleh sebanyak 48 aturan. Penerapan Algoritma C 4.5 dapat melakukan penyederhanaan sebanyak 42 aturan. Aturan yang diperoleh dari pohon keputusan yang didapat adalah 6 aturan.
- d. Hasil analisis perhitungan menunjukkan bahwa terdapat beberapa ranting dari pohon keputusan yang belum mendapatkan keputusan atau class yang disebabkan oleh nilai atribut tujuan tidak ada yang 0 sementara analisis dengan rapidminer langsung diputuskan class dengan mengacu pada nilai yang besar.
- e. Model ini dapat menjadi solusi bagi pengelolaan data dilapangan sehingga membantu dalam pengambilan keputusan terutama berkaitan dengan pencapaian program-program pelayanan yang terencana, terintegrasi dan berkesinambungan.

6.2 Saran

- a. Mengingat dalam penelitian ini terbatas pada data sampel sebanyak 186 sampel, maka untuk penelitian selanjutnya disarankan untuk mengambil populasi seluruhnya disertakan dengan sensus penduduk.
- b. Penelitian ini menitikberatkan pada pemenuhan kebutuhan dasar anak , pada penelitian selanjutnya diharapkan mampu lebih menggali korelasi antara pendapatan, pekerjaan dan tingkat pendidikan terhadap upaya pemenuhan kebutuhan dasar anak.
- c. Meskipun kehandalan algoritma C4.5 terbukti dari penelitian sebelumnya dan dalam penelitian ini, untuk data serupa dalam penelitian ini, bagi penelitian lanjutan sebaiknya dilakukan komparasi kemampuan algoritma klasifikasi yang ada.
- d. Penelitian selanjutnya dapat mengembangkan system berbasis web untuk penentuan keputusan tingkat pemenuhan kebutuhan dasar anak sebagai media sosialisasi atau penyuluhan bisa menggunakan DBMS MYSQL dan JSP atau PHP.

PERNYATAAN ORIGINALITAS

“Saya menyatakan dan bertanggung jawab dengan sebenarnya bahwa Artikel ini adalah hasil karya saya sendiri kecuali cuplikan dan ringkasan yang masing-masing telah saya jelaskan sumbernya”

Yuli Murdianingsih – P31.2010.00810

DAFTAR PUSTAKA

- [1] Hurlock, E. B., Psikologi Anak, Jakarta : Erlangga, 1997.
- [2] UU no 4 tahun 1979 tentang Kesejahteraan Anak
- [3] Fadilah, U., Analisis Penyandang Masalah Kesejahteraan Sosial di Indonesia menggunakan metode FCM Clustering dan Biplot, publikasi ilmiah. [online] 2011.
- [4] Kadushin, A., Child Welfare Services, New York, 1974.
- [5] Chang, C., A Study of Applying *Data mining* to Early Intervention for Developmentally-Delayed Children, ELSEVIER international Jurnal of Expert Systems with Applications 33 (2007) 407–412. [online] 2007.
- [6] Craw, S., Case Base Reasoning : Lecture 3: CBR Case-Base Indexing. [online] 2005. (<http://www.comp.rgu.ac.uk/staff/smc/teaching/kbp3/>).

- [7] Han, J., & Kamber, M., *Data mining* Concept and Tehniques. San Fransisco: Morgan Kauffman, 2006.
- [8] Sayad, S., Model Evaluation-Clasification. [online] 2013. (http://www.saedsayad.com/model_evaluation_c.htm).
- [9] Badan Koordinasi Keluarga Berencana Nasional (BKKBN) Indonesia.
- [10] Prismadia K, Metode Penelitian, FPSI UI, 2008.
- [11] Samian, Skala Likert, 2008.
- [12] Anonymous, Pengolahan data *Training_testing*, 2007.