

IMPLEMENTASI ALGORITMA C4.5 UNTUK KLASIFIKASI ANAK BERKEBUTUHAN KHUSUS DI IBNU SINA STIMULASI CENTER

¹Inayatun Massulloh, ²Fitriyani

¹Universitas Adhirajasa Reswara Sanjaya,
Jl. Sekolah Internasional No. 1-2 Antapani, Telp: 02227100124 Bandung, Jawa Barat
Email: 1inayatunmassulloh59@gmail.com

²Universitas Adhirajasa Reswara Sanjaya,
Jl. Sekolah Internasional No. 1-2 Antapani, Telp: 02227100124 Bandung, Jawa Barat
Email: 2fitriyani@ars.ac.id

ABSTRAK

ABK (Anak Berkebutuhan Khusus) merupakan anak yang secara signifikan berbeda dengan anak-anak lain, mereka yang secara fisik, psikologi, kognitif atau secara sosial terhambat dalam mencapai tujuan dan potensi mereka secara maksimal seperti gangguan bicara, cacat, keterbatasan mental dan gangguan emosional. Klasifikasi anak tunagrahita meliputi anak tunagrahita ringan, sedang dan berat. Tunagrahita meliputi berbagai tingkatan dari yang ringan sampai yang berat. Karena itu mereka berbeda satu dengan yang lain dalam keterlambatan perkembangan dan kondisinya sehingga berbeda dalam strategi pendidikan dan pengajaran yang dirancang dan diprogramkan. Oleh karena itu, diperlukan cara atau metode untuk memberikan informasi kepada guru dan para orang tua dalam membedakan atau mengklasifikasikan anak berkebutuhan khusus (tunagrahita) berdasarkan ciri-ciri yang ada pada anak. Algoritma C4.5 merupakan salah satu metode data mining untuk memprediksi dan mengetahui tingkat akurasi jenis tunagrahita berdasarkan ciri-ciri yang ada pada anak. Dengan menggunakan aplikasi RapidMiner dan pengujian model menggunakan *Cross Validation*, evaluasi dengan *Confusion Matrix* sehingga dihasilkan akurasi dari metode tersebut. Berdasarkan hasil pengujian dengan metode *cross validation* pada aplikasi *RapidMiner* menghasilkan nilai akurasi sebesar 96.18%.

Kata Kunci: Anak Berkebutuhan Khusus, Tunagrahita, Data Mining, Algoritma C4.5, RapidMiner

ABSTRACT

ABK (Children with Special Needs) are children who are significantly different from other children, those who are physically, psychologically, cognitive or socially impeded in achieving their goals and potential to the fullest such as speech disorders, disabilities, mental retardation and disorders emotion. Classifications of mentally retarded children include mild, moderate and severe retarded children. Mental retardation includes various levels from mild to severe. Therefore they differ from each other in the developmental delays and conditions so that they differ in the education and teaching strategies that are designed and programmed. Therefore, we need a way or method to provide information to teachers and parents in differentiating or classifying children with special needs (mental retardation) based on the characteristics that exist in children. C4.5 algorithm is a data mining method to predict and determine the level of accuracy of mental retardation based on the characteristics that exist in children. By using the RapidMiner application and testing the model using Cross Validation, evaluate with Confusion Matrix so that the accuracy of the method is generated. Based on the results of testing with the cross validation method on the RapidMiner application produces an accuracy value of 96.18%.

Keywords: *The Children with Special Needs, mentally disabled, Data Mining, C4.5 Algorithm, RapidMiner*

1. PENDAHULUAN

Pentingnya pendidikan bukan hanya dipahami masyarakat sebagai perwujudan hak warga negara akan pendidikan, tetapi berkembang menjadi suatu kebutuhan bagi pengembangan diri individu, dimana guru diharapkan dapat merealisasikan seluruh potensinya. Hal ini tentu saja, termasuk anak tunagrahita yang tidak dibedakan dalam memperoleh pendidikan walaupun mereka memiliki keterbatasan kecerdasan dan penyesuaian tingkah laku. ABK (Anak Berkebutuhan Khusus) merupakan anak yang secara signifikan berbeda dengan anak-anak lain, mereka yang secara fisik, psikologi, kognitif atau secara sosial terhambat dalam mencapai tujuan dan potensi mereka secara maksimal seperti gangguan bicara, cacat, retardasi mental dan gangguan emosional. Klasifikasi anak tunagrahita meliputi anak tunagrahita ringan, sedang dan berat. Menurut Dudi Gunawan dalam (Garnida, 2016) mengemukakan bahwa “anak gangguan intelektual (Tunagrahita) adalah anak yang secara nyata mengalami hambatan dan keterbelakangan perkembangan mental-intelektual di bawah rata-rata, sehingga mengalami kesulitan dalam menyelesaikan tugas-tugasnya”. Di samping itu, mereka juga mengalami keterbelakangan dalam menyesuaikan diri dengan lingkungan. Pada penelitian ini menggunakan dataset privat sebanyak 106 yang diperoleh dari Ibnu Sina Stimulation Center.

Berdasarkan uraian diatas, bahwa tunagrahita meliputi berbagai tingkatan dari yang ringan sampai yang berat. Karena itu mereka berbeda satu dengan yang lain dalam keterlambatan perkembangan dan kondisinya sehingga berbeda dalam strategi pendidikan dan pengajaran yang dirancang dan diprogramkan.

Ada beberapa penelitian yang telah dilakukan sebelumnya oleh peneliti lain. Penelitian sebelumnya mengenai anak berkebutuhan khusus, akan tetapi menggunakan algoritma yang berbeda seperti penelitian dengan metode (MKNN) Modified K-Nearest Neighbor (Putri et al., 2017), metode business intelligence dengan model Klasifikasi algoritma C4.5 dan Naive Bayes (Grand & Indrajit, 2017). Sedangkan penelitian yang menggunakan algoritma C4.5, akan tetapi menggunakan dataset yang berbeda adalah algoritma C4.5 (Sugara et al., 2018), algoritma Naive Bayes Dan C4.5 (Rahmawati, 2015), algoritma

C4.5 (Jayawardanu, 2015), metode Decision tree (C4.5), Naive Bayes dan k-Nearest Neighbour (Arifin, 2015).

Algoritma C4.5 merupakan salah satu algoritma terbaik dengan urutan paling pertama. Algoritma C4.5 dirancang oleh J. RossQuinlan, dinamakan demikian karena merupakan keturunan dari pendekatan ID3 untuk merangsang pohon keputusan (Fitriyani, 2018). Selain itu, Algoritma C4.5 merupakan salah satu metode data mining untuk mengetahui nilai akurasi klasifikasi dengan menggunakan aplikasi RipedMiner dan metode decision tree C4.5.

Tujuan penelitian ini adalah menerapkan algoritma C4.5 dalam melakukan klasifikasi untuk mengetahui seberapa bagus atau baik hasil prediksi dan hasil nilai akurasi pada data klasifikasi anak berkebutuhan khusus (tunagrahita) menggunakan model *Decision Tree* dan mengimplementasikan algoritma C4.5 untuk mengklasifikasi anak berkebutuhan khusus (tunagrahita) berdasarkan ciri-ciri yang ada pada anak.

2. METODE PENELITIAN

2.1. Analisis Sistem

Peneliti melakukan analisis permasalahan yang terjadi dan melihat kebutuhan apa saja yang dapat menyelesaikan permasalahan.

2.1.1. Analisa Masalah

Masalah pada penelitian ini untuk mengetahui nilai akurasi klasifikasi anak berkebutuhan khusus (tunagrahita) menggunakan algoritma decision tree. Klasifikasi yang dilakukan mengenai klasifikasi anak tunagrahita, yaitu anak tunagrahita ringan, sedang atau berat. Kemudian diukur seberapa baik kinerja model algoritma decision tree pada dataset anak berkebutuhan khusus (tunagrahita) sehingga dapat menghasilkan nilai akurasi yang baik.

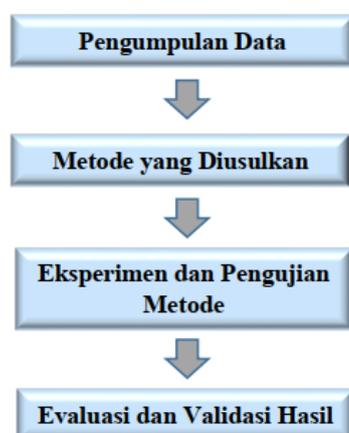
2.1.2. Analisis Sumber Data

Pada penelitian ini, data yang dikumpulkan dan digunakan adalah dataset private sebanyak 106 yang diperoleh dari Ibnu Sina Stimulation Center. Menggunakan dataset private dapat diambil dari organisasi yang kita jadikan obyek.

2.2. Tahapan Penelitian

Tahapan penelitian bertujuan untuk menyelesaikan masalah yang ada secara terstruktur Dalam penelitian ini akan

dilakukan beberapa langkah-langkah atau tahapan penelitian seperti gambar di bawah ini:



Gambar 1. Tahapan penelitian

2.3. Pengumpulan Data

Pada tahap ini peneliti melakukan pengumpulan data. Hal ini merupakan tahapan yang penting karena dapat berpengaruh terhadap hasil penelitian, sehingga dalam mengumpulkan data harus dilakukan dengan benar. Data pada penelitian ini merupakan data primer yang diperoleh dari Ibnu Sina Stimulation Center. Data yang terkumpul sebanyak 106 data.

2.4. Metode yang Diusulkan

Dalam penelitian ini metode yang diusulkan adalah metode klasifikasi data mining algoritma C4.5 dan Decision Tree terhadap data anak berkebutuhan khusus berdasarkan ciri-ciri yang ada pada anak dengan menggunakan RapidMiner. Pengujian model menggunakan Cross Validation, evaluasi dengan Confusion Matrix sehingga dihasilkan akurasi dari metode tersebut.

2.5. Eksperimen dan Pengujian Metode

Pada tahap ini dilakukan eksperimen dan pengujian metode yang digunakan yaitu menghitung dan mendapatkan rule-rule yang ada pada algoritma yang diusulkan yaitu Algoritma C4.5. Eksperimen pada penelitian ini menggunakan laptop LENOVO T420 dengan prosesor Intel Celeron CPU @ 2.50 GHz, memori (RAM) 4.00 GB dan sistem operasi Windows 7 Pro 32-bit serta aplikasi yang digunakan dalam penelitian ini adalah Rapidminer.

2.6. Evaluasi dan Validasi Hasil

Dalam tahap ini dilakukan validasi dan evaluasi untuk mengukur tingkat akurasi data menggunakan beberapa teknik yang terdapat dalam RapidMiner. Validasi dan pengujian dilakukan untuk mengetahui tingkat akurasi, presisi, dan recall dari hasil prediksi klasifikasi. Akurasi adalah persentase dari catatan yang diklasifikasikan dengan benar dalam pengujian dataset. Presisi adalah persentase data yang diklasifikasikan sebagai model yang baik. Recall adalah pengukuran tingkat pengenalan positif sebenarnya.

3. HASIL DAN PEMBAHASAN

3.1. Eksperimen dan Pengujian Metode Algoritma C4.5

Pada tahap ini dilakukan eksperimen dan pengujian metode yang digunakan yaitu menghitung dan menjelaskan basis pengetahuan, yaitu tabel jenis tunagrahita, tabel ciri-ciri tunagrahita, tabel keputusan jenis tunagrahita, rule-rule dan pohon keputusan, yang akan dijelaskan secara rinci sebagai berikut:

3.1.1. Tabel Jenis Tunagrahita

Tabel jenis tunagrahita akan menampilkan jenis tunagrahita yang dialami pada anak.

Kode	Jenis Tunagrahita
J01	Tunagrahita Ringan
J02	Tunagrahita Sedang
J03	Tunagrahita Berat

3.1.2. Tabel Ciri-Ciri Tunagrahita

Tabel ciri-ciri tunagrahita akan menampilkan ciri-ciri tunagrahita yang dialami pada anak.

Atribut	Deskripsi
C01	IQ berkisar antara 50-55 to Aporox, 70
C02	IQ berkisar antara 35-40 to 50-55
C03	IQ berkisar antara 20-25 to 35-40
C04	Sukar mulai dengan sesuatu
C05	Termasuk kelompok mampu didik
C06	Termasuk kelompok mampu rawat
C07	Mengerjakan sesuatu berulang-ulang
C08	Kemampuan motorik yang kurang

C09	Perhatian yang mudah teralih
C10	Kesulitan menyesuaikan diri
C11	Adanya kesulitan belajar
C12	Termasuk kelompok mampu latih
C13	Mempunyai kemampuan untuk berkembang dalam bidang pelajaran akademik, penyesuaian sosial dan kemampuan bekerja
C14	Mampu melakukan keterampilan mengurus dirinya sendiri
C15	Dapat mandiri
C16	Masih dapat mengikuti pekerjaan baik di sekolah maupun di sekolah umum
C17	Kurangnya perbendaharaan kata
C18	Kelemahan artikulasi
C19	Pemusatan perhatian sedikit hyperactive
C20	Sepanjang hidupnya tergantung pada bantuan orang lain
C21	Karakteristik fisik Nampak seperti anak normal hanya sedikit mengalami kelemahan dalam kemampuan sensomotorik
C22	Dapat melakukan pekerjaan yang semi terampil, atas pekerjaan tertentu yang dapat dijadikan bekal bagi hidupnya.
C23	Kemampuan belajar paling tinggi setaraf anak normal usia 7 atau 8 tahun
C24	Kemampuan belajar paling tinggi setaraf anak normal usia 3-4 tahun
C25	Kemampuan belajar paling tinggi setaraf anak normal usia 12 tahun

3.1.3. Tabel Keputusan Jenis Tunagrahita

Tabel keputusan jenis tunagrahita ini menampilkan relasi antara ciri-ciri tunagrahita dengan jenis tunagrahita yang dialami oleh anak.

Tabel 3.
Tabel Keputusan Jenis Tunagrahita

Atribut	Jenis Tunagrahita		
	J01	J02	J03
C01	Ya	Tidak	Tidak
C02	Tidak	Ya	Tidak
C03	Tidak	Tidak	Ya
C04	Ya	Ya	Ya
C05	Ya	Tidak	Tidak
C06	Ya	Ya	Ya
C07	Ya	Ya	Ya

C08	Tidak	Ya	Ya
C09	Ya	Ya	Ya
C10	Ya	Ya	Ya
C11	Ya	Ya	Ya
C12	Ya	Ya	Tidak
C13	Ya	Tidak	Tidak
C14	Ya	Tidak	Tidak
C15	Ya	Tidak	Tidak
C16	Ya	Tidak	Tidak
C17	Ya	Ya	Ya
C18	Tidak	Ya	Ya
C19	Tidak	Tidak	Ya
C20	Tidak	Tidak	Ya
C21	Tidak	Tidak	Tidak
C22	Ya	Ya	Tidak
C23	Tidak	Ya	Tidak
C24	Tidak	Tidak	Ya
C25	Ya	Tidak	Tidak

3.2. Analisa Perhitungan

Pada tahap ini dilakukan eksperimen dan pengujian metode yang digunakan yaitu menghitung dan mendapatkan rule-rule yang ada pada algoritma yang diusulkan yaitu Algoritma C4.5 Menurut Kusriani dan Lutfi dalam (Haryati et al., 2015) Ada beberapa tahap dalam membuat sebuah pohon keputusan dengan algoritma C4.5, yaitu:

1. Menyiapkan data training. Data training biasanya dari data histori yang pernah terjadi sebelumnya dan sudah dikelompokkan ke dalam kelas-kelas tertentu.
2. 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.
3. Sebelum nilai *Gain* dihitung, nilai entropy akan dihitung terlebih dahulu
4. Setelah itu, hitung kembali dengan cara yang sama untuk semua simpul yang terpartisi.
5. Proses partisi pohon keputusan akan berhenti saat:
 - a. Semua tupel dalam node N mendapat kelas yang sama.
 - b. Tidak ada atribut di dalam tupel yang dipartisi lagi.
 - c. Tidak ada tupel di dalam cabang yang kosong

Tabel 4. Tabel Analisa Perhitungan Nilai Entropy dan Gain

Ciri-ciri Tunagrahita	Jumlah Kasus	Tunagrahita Ringan	Tunagrahita Sedang	Tunagrahita Berat	Entropy	Gain
Total	106	74	26	6	1,093762372	
C01						0,753798443
Ya	74	73	1	0	0	
Tidak	32	1	25	6	0,887306883	
C02						0,673929281
Ya	26	1	25	0	0,235193382	
Tidak	80	73	1	6	0,479840997	
C03						0,313812964
Ya	6	0	0	6	0	
Tidak	100	74	26	0	0,826746372	
C04						0,025177801
Ya	101	69	26	6	1,121484798	
Tidak	5	5	0	0	0	
C05						0,753798443
Ya	74	73	1	0	0,103275085	
Tidak	32	1	25	6	0,887306883	
C06						0
Ya	106	74	26	6	1,093762372	
Tidak	0	0	0	0	0	
C07						0,025066939
Ya	101	70	26	5	1,08523243	
Tidak	5	4	0	1	0,721928095	
C08						0,095779677
Ya	81	50	25	6	1,231220506	
Tidak	25	24	1	0	0,242292189	
C09						0,164929208
Ya	78	46	26	6	1,262260454	
Tidak	28	28	0	0	0	
C10						0
Ya	106	74	26	6	1,093762372	
Tidak	0	0	0	0	0	

Dari tabel analisa perhitungan diatas, akan diperhitungkan nilai *entropy* dan *gain* dari masing-masing jenis tunagrahita.

- Menghitung jumlah kasus “tunagrahita ringan”, “tunagrahita sedang” dan “tunagrahita berat” serta nilai *entropy* dari semua kasus. Dari data training yang ada, diketahui jumlah kasus “tunagrahita ringan” sebanyak 74 *record*, jumlah kasus “tunagrahita sedang” sebanyak 26 *record* dan jumlah kasus “tunagrahita berat” sebanyak 6 *record*. Total kasus keseluruhan adalah 106 kasus, sehingga di dapat *entropy* keseluruhan:

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

$$\begin{aligned} Entropy (Total) &= ((-74/106)*\log_2(74/106) + \\ &(-26/106)*\log_2(26/106)+ \\ &(6/106)*\log_2(6/106)) \\ &= 1,093762372 \end{aligned}$$

- Hitung nilai *entropy* dan nilai *gain* masing-masing atribut. Nilai *gain* tertinggi adalah atribut yang menjadi *root* dari pohon keputusan yang akan dibuat, dengan rumus sebagai berikut:

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

- C01

$$\begin{aligned} Entropy (Ya) &= ((-73/74)*\log_2(73/74) + \\ &(-1/74)*\log_2(1/74) + \\ &(-0/74)*\log_2(0/74)) \\ &= 0 \end{aligned}$$

$$\begin{aligned} Entropy (tidak) &= ((-1/32)*\log_2(1/32) + \\ &(-25/32)*\log_2(25/32) + \\ &(-6/32)*\log_2(6/32)) \\ &= 0,887306883 \end{aligned}$$

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

$$\begin{aligned} Gain (Total C01) &= 1,093762372 - (74/106)*0) - \\ &((32/106)* 0,887306883) \\ &= 0,753798443 \end{aligned}$$

- C02

$$\begin{aligned} Entropy (Ya) &= ((-1/26)*\log_2(1/26) + \\ &(-25/26)*\log_2(25/26) + \\ &(-0/26)*\log_2(0/26)) \\ &= 0,235193382 \end{aligned}$$

$$\begin{aligned} Entropy (tidak) &= ((-73/80)*\log_2(73/80) + \\ &(-1/80)*\log_2(1/80) + \\ &(-6/80)*\log_2(6/80)) \\ &= 0,479840997 \\ Gain (Total C02) &= 1,093762372 - \\ &(26/106)*0,235193382) - \\ &((80/106)* 0,479840997) \\ &= 0,673929281 \end{aligned}$$

- C03

$$\begin{aligned} Entropy (Ya) &= ((-0/6)*\log_2(0/6) + \\ &(-0/6)*\log_2(0/6)+ \\ &(-6/6)*\log_2(6/6)) \\ &= 0 \end{aligned}$$

$$\begin{aligned} Entropy (Tidak) &= ((-74/100)*\log_2(74/100) + \\ &(-26/100)*\log_2(26/100) + \\ &(-0/100)*\log_2(0/100)) \\ &= 0,826746372 \end{aligned}$$

$$\begin{aligned} Gain (Total C03) &= 1,093762372 - ((6/106)*0) - \\ &((100/106)*0,826746372) \\ &= 0,313812964 \end{aligned}$$

- C04

$$\begin{aligned} Entropy (Ya) &= ((-69/101)*\log_2(69/101) + \\ &(-26/101)*\log_2(26/101) + \\ &(-6/101)*\log_2(6/101)) \\ &= 1,121484798 \end{aligned}$$

$$\begin{aligned}
 & \text{Entropy (Tidak)} \\
 & = ((-5/5) \cdot \log_2(5/5) + \\
 & \quad (-0/5) \cdot \log_2(0/5) + \\
 & \quad (-0/5) \cdot \log_2(0/5)) \\
 & = 0
 \end{aligned}$$

$$\begin{aligned}
 & \text{Gain (Total C04)} \\
 & = 1,093762372 - \\
 & \quad ((101/106) \cdot 1,121484798) - \\
 & \quad ((5/106) \cdot 0) \\
 & = 0,025177801
 \end{aligned}$$

e. CO5

$$\begin{aligned}
 & \text{Entropy (Ya)} \\
 & = ((-73/74) \cdot \log_2(73/74) + \\
 & \quad (-1/74) \cdot \log_2(1/74) + \\
 & \quad (-0/74) \cdot \log_2(0/74)) \\
 & = 0,103275085
 \end{aligned}$$

$$\begin{aligned}
 & \text{Entropy (tidak)} \\
 & = ((-1/32) \cdot \log_2(1/32) + \\
 & \quad (-25/32) \cdot \log_2(25/32) + \\
 & \quad (-6/32) \cdot \log_2(6/32)) \\
 & = 0,887306883
 \end{aligned}$$

$$\begin{aligned}
 & \text{Gain (Total C05)} \\
 & = 1,093762372 - \\
 & \quad ((74/106) \cdot 0,103275085) - \\
 & \quad ((32/106) \cdot 0,887306883) \\
 & = 0,753798443
 \end{aligned}$$

f. CO6

$$\begin{aligned}
 & \text{Entropy (Ya)} \\
 & = ((-74/106) \cdot \log_2(74/106) + \\
 & \quad (-26/106) \cdot \log_2(26/106) + \\
 & \quad (-6/106) \cdot \log_2(6/106)) \\
 & = 1,093762372
 \end{aligned}$$

$$\text{Entropy (tidak)} = 0$$

$$\begin{aligned}
 & \text{Gain (Total C06)} \\
 & = 1,093762372 - \\
 & \quad ((106/106) \cdot 1,093762372) - \\
 & \quad ((0/106) \cdot 0) \\
 & = 0
 \end{aligned}$$

g. CO7

$$\begin{aligned}
 & \text{Entropy (Ya)} \\
 & = ((-70/101) \cdot \log_2(70/101) + \\
 & \quad (-26/101) \cdot \log_2(26/101) + \\
 & \quad (-5/101) \cdot \log_2(5/101)) \\
 & = 1,08523243
 \end{aligned}$$

$$\begin{aligned}
 & \text{Entropy (tidak)} \\
 & = ((-4/5) \cdot \log_2(4/5) + (-0/5) \cdot \log_2(0/5) + \\
 & \quad (-1/5) \cdot \log_2(1/5)) \\
 & = 0,721928095
 \end{aligned}$$

$$\begin{aligned}
 & \text{Gain (Total C07)} \\
 & = 1,093762372 - \\
 & \quad ((101/106) \cdot 1,08523243) - \\
 & \quad ((5/106) \cdot 0,721928095) \\
 & = 0,025666939
 \end{aligned}$$

h. CO8

$$\begin{aligned}
 & \text{Entropy (Ya)} \\
 & = ((-50/81) \cdot \log_2(50/81) + \\
 & \quad (-25/81) \cdot \log_2(25/81) + \\
 & \quad (-6/81) \cdot \log_2(6/81)) \\
 & = 1,231220506
 \end{aligned}$$

$$\begin{aligned}
 & \text{Entropy (tidak)} \\
 & = ((-24/25) \cdot \log_2(24/25) + \\
 & \quad (-1/25) \cdot \log_2(1/25) + \\
 & \quad (-0/25) \cdot \log_2(0/25)) \\
 & = 0,242292189
 \end{aligned}$$

$$\begin{aligned}
 & \text{Gain (Total C08)} \\
 & = 1,093762372 - \\
 & \quad ((81/106) \cdot 1,231220506) - \\
 & \quad ((25/106) \cdot 0,242292189) \\
 & = 0,095779677
 \end{aligned}$$

i. CO9

$$\begin{aligned}
 & \text{Entropy (Ya)} \\
 & = ((-46/78) \cdot \log_2(46/78) + \\
 & \quad (-26/78) \cdot \log_2(26/78) + \\
 & \quad (-6/78) \cdot \log_2(6/78)) \\
 & = 1,262260454
 \end{aligned}$$

$$\begin{aligned}
 & \text{Entropy (tidak)} \\
 & = ((-28/28) \cdot \log_2(28/28) + \\
 & \quad (-0/28) \cdot \log_2(0/28) + \\
 & \quad (-0/28) \cdot \log_2(0/28)) \\
 & = 0
 \end{aligned}$$

$$\begin{aligned}
 & \text{Gain (Total C09)} \\
 & = 1,093762372 - \\
 & \quad ((78/106) \cdot 1,262260454) - \\
 & \quad ((28/106) \cdot 0) \\
 & = 0,164929208
 \end{aligned}$$

j. C10

$$\begin{aligned}
 & \text{Entropy (Ya)} \\
 & = ((-74/106) \cdot \log_2(74/106) + \\
 & \quad (-26/106) \cdot \log_2(26/106) + \\
 & \quad (-6/106) \cdot \log_2(6/106)) \\
 & = 1,093762372
 \end{aligned}$$

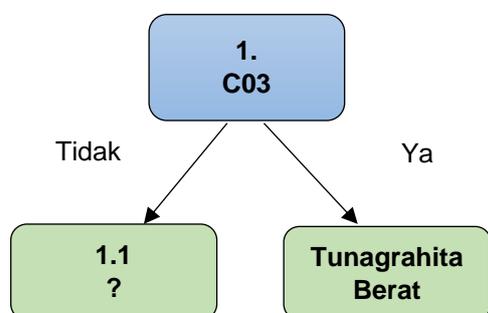
$$\text{Entropy (tidak)} = 0$$

$$\begin{aligned}
 & \text{Gain (Total C10)} \\
 & = 1,093762372 - \\
 & \quad ((106/106) \cdot 1,093762372) - \\
 & \quad ((0/106) \cdot 0) \\
 & = 0
 \end{aligned}$$

Lakukan perhitungan dengan cara yang sama bagi atribut C11 – C25 untuk mendapat nilai *entropy* dan *gain*. Berdasarkan hasil perhitungan nilai *entropy* dan *gain* yang terdapat pada tabel 4, bahwa nilai *gain* terbesar adalah C03, yakni sebesar 0,313812964, sehingga atribut C03 digunakan sebagai pohon atau root, kemudian buat cabang sesuai banyaknya partisi. Pada atribut C03 terdapat dua partisi yakni “Ya” dan “Tidak”, sehingga dari root C03 akan dibentuk dua cabang “Ya” dan “Tidak”.

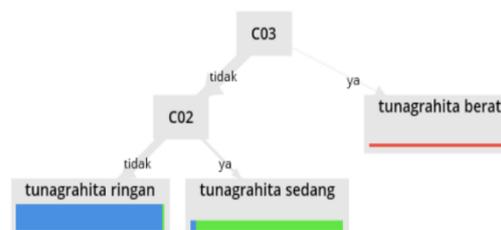
C03	Jumlah Kasus	Tunagrahita Ringan	Tunagrahita Sedang	Tunagrahita Berat
Ya	6	0	0	6
Tidak	100	74	26	0

Dari masing-masing cabang, perhatikan partisi yang memiliki kasus sama dengan 0. Dari data diatas, diketahui bahwa partisi “Ya” memiliki nilai 0 pada kasus Tunagrahita ringan dan tunagrahita sedang, kemudian pada atribut “C03”, memiliki 6 kasus dan semuanya memiliki jawaban Ya ($\text{Sum}(\text{Total}) / \text{Sum}(\text{Ya}) = 6/6 = 1$). Dengan demikian C03 “Ya” menjadi daun atau leaf, sehingga untuk C03 “Ya” akan menghasilkan label “Tunagrahita Berat”, maka pohon keputusan dapat dibuat sebagai berikut:



Berdasarkan pembentukan pohon keputusan node 1 (root node), node (simpul) 1.1 akan dilakukan kembali perhitungan nilai *entropy* dan *gain* atau akan dianalisis lebih lanjut dengan cara yang sama.

Kemudian untuk mencari simpul dari partisi cabang-cabang selanjutnya, hitung kembali nilai *entropy* dan nilai *gain* dari setiap atribut. Setiap atribut yang memiliki nilai *gain* terbesar atau tertinggi akan menjadi daun pohon atau leaf selanjutnya. Setelah dilakukan perhitungan secara keseluruhan, maka diperoleh pohon keputusan seperti gambar berikut ini:



Gambar 2.

Pohon keputusan (*Decision Tree*)

Setelah pohon keputusan (*decision tree*) terbentuk, adapun aturan atau rule yang terbentuk dari pohon keputusan (*decision tree*) terakhir berdasarkan pada gambar 4.2 adalah:

Rule 1

IF C03 == Ya
THEN Tunagrahita Berat

Rule 2

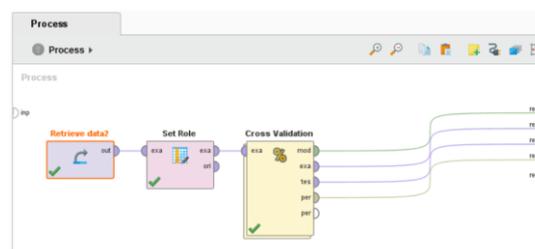
IF C03 == Tidak
AND C02 == Ya
THEN Tunagrahita Sedang

Rule 2

IF C03 == Tidak
AND C02 == Tidak
THEN Tunagrahita Ringan

3.3. Pengujian Metode Algoritma C4.5

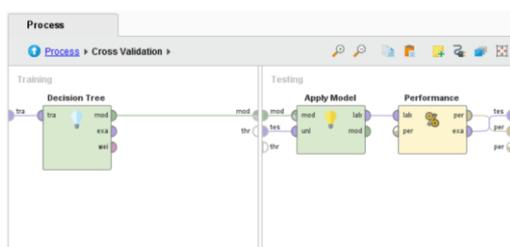
Pengujian dengan menggunakan cross validation pada aplikasi RapidMiner dapat dilihat pada gambar dibawah ini:



Gambar 3. Pengujian *cross validation*

Model yang telah dibentuk diuji tingkat akurasi dengan memasukkan atau uji yang berasal dari data training dengan modeling “*Tree – Decision Tree*”, lalu

testing dengan *Apply Model & Validation* "Performance – Predictive %Performance (Clasification) pada aplikasi *rapidMiner* untuk menguji tingkat akurasi.



Gambar 4. Validation model algoritma C4.5

3.4. Evaluasi dan Validasi Hasil

Setelah data diolah maka dapat diuji tingkat akurasinya untuk melihat kinerja dari metode C4.5. Penelitian ini bertujuan untuk melihat akurasi analisis data anak berkebutuhan khusus (Tunagrahita), menilai kemungkinan ciri-ciri yang ada pada anak apakah termasuk ke dalam anak berkebutuhan khusus jenis tunagrahita ringan, sedang atau berat. Pengujian tingkat akurasi dilakukan dengan menggunakan confusion matrix.

accuracy: 96.18% +/- 6.54% (micro average: 96.23%)

	true tunagrahita ringan	true tunagrahita sedang	true tunagrahita berat	class precision
pred. tunagrahita ringan	73	3	0	96.05%
pred. tunagrahita sedang	1	23	0	95.83%
pred. tunagrahita berat	0	0	5	100.00%
class recall	98.65%	88.46%	100.00%	

Gambar 5. Nilai akurasi dari confusion matrix algoritma c4.5

Dari gambar di atas dijelaskan bahwa tunagrahita ringan menghasilkan *class recall* sebesar 98.65% dan *class precision* sebesar 96.05%, sedangkan tunagrahita sedang menghasilkan *class recall* sebesar 88.46% dan *class precision* sebesar 95.83% serta tunagrahita berat menghasilkan *class recall* sebesar 100% dan *class precision* sebesar 100% dan nilai akurasi dari perhitungan algoritma C4.5 tersebut yaitu sebesar 96.18%.

4. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan pada penerapan klasifikasi decision tree dengan algoritma C4.5 maka dapat disimpulkan bahwa:

- Klasifikasi anak berkebutuhan khusus dengan memanfaatkan teknik data mining menggunakan metode algoritma C4.5 dapat memprediksi jenis tunagrahita berdasarkan ciri-ciri yang

ada pada anak. Setelah dilakukan pengujian dengan metode *cross validation* pada aplikasi *RapidMiner* menghasilkan nilai akurasi sebesar 96.18%.

- Dengan adanya penerapan data mining algoritma C4.5 diharapkan mampu memberikan solusi dalam menentukan anak berkebutuhan khusus jenis tunagrahita di Ibnu Sina Stimulation Center.

REFERENSI

- Arifin, T. (2015). *METODE DATA MINING UNTUK KLASIFIKASI DATA SEL NUKLEUS DAN SEL RADANG BERDASARKAN ANALISA TEKSTUR*. II(2), 425–433.
- Desiningrum, D. R. (2016). *Psikologi anak berkebutuhan khusus*.
- Fitriyani, F. (2016). Implementasi Algoritma Fp-Growth Menggunakan Association Rule Pada Market Basket Analysis. *Jurnal Informatika*, 2(1). <https://doi.org/10.31311/ji.v2i1.85>
- Fitriyani, F. (2018). Metode Bagging Untuk Imbalance Class Pada Bedah Toraks Menggunakan Naive Bayes. *Jurnal Kajian Ilmiah*, 18(3), 278. <https://doi.org/10.31599/jki.v18i3.281>
- Garnida, D. (2016). Modul Guru Pembelajaran SLB TUNAGRAHITA. *Pppptk Tk Dan Plb Bandung*, 1–165.
- Grand, & Indrajit, R. E. (2017). Aplikasi Deteksi Dini Anak Berkebutuhan Khusus Menggunakan Metode Business. *Seminar Nasional Sains Dan Teknologi 2017, November*, 1–10.
- Jayawardanu, I. H. and S. H. (2015). *Rancang Bangun Sistem Pakar untuk Deteksi*. VII(2), 48–58.
- Putri, Z. S., Putri, R. R. M., Indriati, Hariyanto, E., Witanti, A., Zamroni, I., Werdiningsih, I., Purbandini, Budiman, E., Santoso, E., Afirianto, T., Susanto, I. D., Rifai, B., Kuswanto, H., Rahmawati, A., Wintana, D., Suhada, S., Gunawan, Sulaiman, H., ... Nurcahyo, G. W. (2017). Deteksi

-
- Autisme pada Anak Menggunakan Metode Modified K-Nearest Neighbor (MKNN). *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 1(2), 37–43. <https://doi.org/10.30873/ji.v15i2.598>
- Rahmawati, E. (2015). Analisa Komparasi Algoritma Naive Bayes Dan C4.5 Untuk Prediksi Penyakit Liver. *None*, 12(2), 27–37.
- Sugara, B., Widyatmoko, D., Prakoso, B. S., & Saputro, D. M. (2018). Penerapan Algoritma C4.5 untuk Deteksi Dini Autisme Pada Anak. *Seminar Nasional Teknologi Informasi Dan Komunikasi (SENTIKA), 2018*(Sentika), 87–96.
- Zaka, Q. (2016). *DALAM MENGIKUTI PEMBELAJARAN PENDIDIKAN JASMANI DI SLB-C YAYASAN PENDIDIKAN LUAR BIASA DEMAK TAHUN 2016 SKRIPSI UNIVERSITAS NEGERI SEMARANG TAHUN 2016 ABSTRAK.*