

SISTEM PENDUKUNG KEPUTUSAN PENERIMAAN BEASISWA MENGUNAKAN ALGORITMA C4.5 (STUDI KASUS : BAITUL MAAL)

Agustiena Merdekawati

Manajemen Informatika, AMIK Bina Sarana Informatika
email: agustiena.atd@bsi.ac.id

Abstrak

Beasiswa merupakan sebuah cara untuk mengatasi masalah biaya atau ekonomi bagi mereka yang kurang mampu. Beasiswa adalah pemberian berupa bantuan keuangan yang diberikan kepada perorangan yang bertujuan untuk digunakan demi keberlangsungan pendidikan yang ditempuh. Pemberi beasiswa, yaitu lembaga pemerintahan, perusahaan, atau yayasan. Pemberian beasiswa ini sangat penting, guna memberikan ketenangan mereka dalam mengikuti belajar-mengajar. Namun terkadang pemberian beasiswa tidak pada orang yang seharusnya menerima atau tidak tepat pada sasarannya. Dari permasalahan tersebut dilakukan penelitian mengenai seleksi penerimaan beasiswa dengan menggunakan algoritma C4.5. Setelah dilakukan pengujian, maka hasil yang didapat dari algoritma C4.5 memiliki akurasi sebesar 69.73%. Dengan adanya penelitian diharapkan dapat membantu baitul maal atau pemberi beasiswa dalam menentukan kriteria yang lebih diutamakan untuk pemberian beasiswa, sehingga pemberian beasiswa dapat tepat sasaran.

Kata kunci: Pendukung keputusan, Beasiswa, Algoritma C4.5

Abstract

Scholarships are a way to address cost or economic issues for the underprivileged. The scholarship is a grant of financial assistance given to individuals who aim to be used for the sustainability of education pursued. Scholarship grantee, ie government institution, company, or foundation. The provision of scholarships is very important, in order to provide their calm in following the teaching and learning. But sometimes scholarships are not on people who should receive or not right on the target. From the problem is done research about the selection of scholarship acceptance by using algorithm C4.5. After testing, the results obtained from the algorithm C4.5 has an accuracy of 69.73%. With the research is expected to help baitul maal or scholarship in determining the criteria that are preferred for the provision of scholarships, so that the scholarship can be right on target.

Keywords: Decision support, Scholarships, Algorithms C4.5

PENDAHULUAN

Seperti yang dituangkan dalam Undang-Undang Dasar 1945 pasal 31 (1) bahwa tiap warga negara berhak mendapatkan pengajaran. Akan tetapi untuk mendapatkan pendidikan itu semakin tidak terjangkau atau sangatlah mahal, terutama bagi yang bertingkat ekonomi menengah kebawah, baik itu tingkat SD, SMP, SMA, bahkan perguruan tinggi.

Beasiswa merupakan salah satu solusi untuk mengatasi masalah biaya bagi mereka yang kurang mampu.

Menurut Murniasih dalam Rainer (2017) beasiswa diartikan "sebagai bentuk penghargaan yang diberikan kepada individu agar dapat melanjutkan pendidikan kejenjang yang lebih tinggi".

Sehingga dengan adanya beasiswa dapat menciptakan pengaruh yang positif bagi mereka, yaitu:

1. Mereka tidak perlu lagi khawatir dengan tagihan biaya pendidikan,
2. Anak tersebut berusaha untuk mendapatkan nilai dan prestasi yang tinggi.

Dalam Penelitian Hijriana dan Rasyidan (2017), dengan kesimpulan Atribut yang berpengaruh terhadap proses seleksi calon penerima beasiswa adalah Indeks Prestasi Semester, Penghasilan yang dibuat menjadi 3 atau 4 kategori kelas (Klasifikasi Penghasilan) dan Tanggungan, baik dengan 2 kategori kelas maupun dengan 3 kategori kelas (Klasifikasi Tanggungan). Sehingga data atribut Penghasilan dan Tanggungan akan lebih berpengaruh dalam penentuan nilai akurasi jika diubah ke dalam bentuk klasifikasi terlebih dahulu dibandingkan jika hanya dengan data asli berupa angka. Hasil evaluasi dengan tool Confussion Matrix yang ditunjukkan melalui tabel Performance Vector C4.5 dan validasi dengan teknik Split Validation menunjukkan tingkat akurasi tertinggi didapat pada eksperimen ke-10 yakni sebesar 85.44%, dengan nilai class precision berturut-turut sebesar 87.61% dan 77.66% dan nilai class recall sebesar 93.54% dan 62.93%.

Berdasarkan kesimpulan dalam penelitian Praja dan Astuti (2016) adalah berdasarkan hasil pengujian metode pohon keputusan terhadap data siswa dapat disimpulkan bahwa, Penerapan metode pohon keputusan terhadap data siswa SMA N 1 Mlonggo memiliki tingkat akurasi yang cukup baik dalam menyelesaikan klasifikasi rekomendasi beasiswa. Metode pohon keputusan merupakan metode yang cukup sesuai untuk penyelesaian studi kasus dalam pemilihan siswa yang mendapatkan rekomendasi beasiswa. Tingkat akurasi yang dihasilkan oleh metode tersebut adalah 79%.

Dalam penelitian Gutterres, beasiswa merupakan sarana yang diberikan dan diajukan oleh institusi pendidikan guna membantu para mahasiswa yang kurang mampu namun memiliki segudang prestasi dalam bidang akademik. Untuk menentukan Mahasiswa layak atau tidak menerima beasiswa maka diperlukan sebuah sistem yang dapat memberikan kemudahan dalam pengajuan penerima beasiswa. Tujuan dari penelitian ini untuk menguji kelayakan algoritma C4.5 sebagai pendukung keputusan dalam pengajuan

penerima beasiswa pada STIKOM Artha Buana. Metode yang digunakan adalah metode klasifikasi menggunakan algoritma C45 dengan kriteria antara lain Tidak memiliki kendaraan pribadi, jarak tempat tinggal jauh \pm 2Km dari kampus, IPK minimal 2,75, belum pernah mendapatkan beasiswa dan penghasilan orang tua < 1 juta. Hasil dari penelitian adalah menentukan mahasiswa sebagai calon yang seharusnya diajukan oleh kampus STIKIM ARTHA BUANA Kupang untuk menerima beasiswa pendidikan berdasarkan hasil perhitungan algoritma C45. Dari hasil perbandingan metode, maka dapat diambil kesimpulan bahwa algoritma C45 layak dijadikan sebagai pendukung keputusan dalam pengajuan penerima beasiswa pada Stikom Artah Buana kupang karena hasil persentasi dari algoritma C45 lebih tinggi daripada metode klasifikasi yang lainnya dan memiliki kendaraan pribadi juga layak dijadikan salah satu kriteria dalam pengajuan penerima dana bantuan beasiswa dari pemerintah daerah ataupun dari dikti/kopertis wilayah 8.

Akan tetapi, pengambilan keputusan masih terdapat beberapa permasalahan yang harus dicarikan solusi terbaik, yang berisi elemen-elemen atau hubungan-hubungan antar elemen yang tidak dapat dipahami dimana masalah tak terstruktur terjadi.

Sedangkan masalah semi terstruktur adalah masalah yang berisi sebagian elemen-elemen atau hubungan yang dimengerti oleh pemecah masalah.

Menurut Hasan SPK dirancang untuk mendukung seluruh tahap pengambilan keputusan mulai dari mengidentifikasi masalah, memilih data yang relevan, menentukan pendekatan yang digunakan dalam proses pengambilan keputusan sampai mengevaluasi pemilihan alternative. Sistem pendukung keputusan (SPK) adalah sebuah sistem yang dimaksudkan untuk mendukung para pengambilan keputusan semi terstruktur. Sistem Pendukung Keputusan (SPK) yang dimaksud untuk menjadi alat bantu bagi para pengambil keputusan untuk memperluas kapasitas, namun tidak untuk

menggantikan penilaian. (Ahda dan Bahri, 2017)

Pada dasarnya pengambilan keputusan adalah suatu pendekatan sistematis pada hakekat suatu masalah, pengumpulan fakta-fakta, penentuan yang matang dari alternatif yang dihadapi dan pengambilan tindakan yang menurut perhitungan merupakan tindakan yang paling tepat. Sekalipun didukung oleh potensi ekosistem dan aksesibilitas yang serba prospektif, pengambilan keputusan kerap dihadapkan pada masalah utama dalam penentuan keputusan strategis yang sulit direalisasikan akibat persepsi yang heterogen sejalan dengan kepentingan masing-masing individu/kelompok yang terlibat dalam pengambilan keputusan. (Suryadi, 2007)

Setiap tahunnya yayasan baitul maal menyediakan beasiswa bagi siswa atau siswi di Indonesia yang kurang mampu dan berprestasi. Akan tetapi tidak semua pendaftar mendapatkan beasiswa. Sistem seleksi beasiswa yang tidak akurat menyebabkan beasiswa salah sasaran.

Penelitian mengenai analisis kelayakan pemberian beasiswa dengan metode klasifikasi data mining telah banyak dilakukan, namun sangat jarang sekali yang menganalisis kelayakan pemberian beasiswa untuk SD, SMP, SMA, dan dengan menggunakan pengujian tingkat keakuratan dengan menggunakan Rapidminer.

Manfaat penelitian ini adalah:

1. Manfaat Praktisi, hasil penelitian ini diharapkan agar dapat digunakan oleh baitul maal dalam penentuan seleksi pemberian beasiswa bagi SD, SMP, dan SMA.
2. Manfaat teoritis, hasil penelitian ini diharapkan dapat memberikan sumbangan bagi penelitian yang berkaitan dengan klasifikasi data mining, khususnya dalam penentuan seleksi pemberian beasiswa bagi SD, SMP, dan SMA.
3. Manfaat kebijakan, dari penelitian ini diharapkan dapat memberikan hasil yang lebih akurat sehingga memudahkan bagian staff baitul maal untuk menyeleksi proses penerimaan beasiswa.

Menurut hermawati (2013:3) Data mining proses yang mempekerjakan satu atau lebih teknik pembelajaran komputer (*machine learning*) untuk menganalisis dan mengekstraksi pengetahuan (*knowledge*) secara otomatis. Definisi lain diantaranya adalah pembelajaran berbasis (*induction-based learning*) adalah proses pembentukan definisi-definisi konsep umum yang dilakukan dengan caramengobservasi contoh-contoh spesifik dari konsep-konsep yang akan dipelajari. *Knowledge Discovery in Databases* (KDD) adalah penerapan metode saintik pada data mining. Dalam konteks ini data mining merupakan satu langkah dari proses KDD.

Menurut Hermawati (2013:3) Data mining merupakan proses iterative dan interaktif untuk menemukan pola atau model baru yang sah (sempurna), bermanfaat dan dapat dimengerti dalam suatu database yang sangat besar (*massive databases*).

- a. Sahih
Dapat digeneralisasi untuk masa yang akan datang.
- b. Baru
Apa yang sedang tidak diketahui.
- c. Bermanfaat
Dapat digunakan untuk melakukan suatu tindakan.
- d. Iteratif
memerlukan sejumlah proses yang diulang
- e. Interaktif
f. memerlukan interaksi manusia dalam prosesnya.

Menurut Hermawati (2013:14) teknik data mining dibagi menjadi beberapa teknik. Beberapa teknik dan sifat data mining adalah sebagai berikut:

- a. Klasifikasi (*Classification*)
Menentukan sebuah record data baru kesalah satu dari beberapa katagori (atau klas) yang telah didefinisikan sebelumnya.
- b. Regresi (*Regression*)
Memprediksi nilai dari suatu variabel kontinyu yang diberikan berdasarkan nilai dari variabel yang lain, dengan mengasumsikan sebuah model ketergantungan linier atau non linier.

Teknik ini banyak dipelajari dalam statistika, bidang jaringan saraf tiruan (neural network).

c. Klasterisasi (*Clustering*)

Mempartisi data-set menjadi beberapa sub-set atau kelompok sedemikian rupa sehingga elemen-elemen dari suatu kelompok tertentu memiliki set properti yang di share bersama, dengan tingkat similitas yang tinggi dalam satu kelompok dan tingkat similitas antar kelompok yang rendah. Disebut juga dengan 'unsupervised learning'.

d. Asosiasi (*Association Rules*)

Mendeteksi sekumpulan atribut-atribut yang muncul bersamaan (co-occur) dalam frekuensi yang sering dan membentuk sejumlah kaidah dari kumpulan-kumpulan tersebut.

e. Sekuensial (*Sequence Mining*)

Mencari sejumlah event yang secara umum terjadi bersama-sama.

Menurut Hermawati (2013:58) Pohon keputusan adalah sebuah struktur data yang terdiri dari simpul (*node*), dan rusuk (*edge*). Simpul pada sebuah pohon dibedakan menjadi tiga, yaitu simpul akar (*root node*), simpul percabangan atau internal (*branch/internal node*) dan simpul daun (*leaf node*).

Pohon keputusan merupakan representasi sederhana dari teknik klasifikasi untuk sejumlah kelas berhingga, dimana simpul internal maupun simpul akar ditandai dengan nama atribut, rusuk-rusuknya diberi label nilai atribut yang mungkin dan simpul daun ditandai dengan kelas-kelas yang berbeda.

Pohon keputusan dibangun berdasarkan suatu algoritma induksi. Salah satu algoritma yang digunakan untuk membangun pohon keputusan yang berbasis algoritma induksi pohon keputusan seperti ID3, C4.5, dan CART adalah algoritma Hunt.

Menurut Hermawati (2013:60) Algoritma C4.5 merupakan algoritma yang digunakan untuk membangun pohon keputusan.

Menurut Kusriani dan Luthfi dalam Haryati,dkk (2015) Ada beberapa tahap dalam membuat sebuah pohon keputusan dengan algoritma C4.5, yaitu :

1. Menyiapkan data training,

Data training biasanya diambil dari data histori yang pernah terjadi sebelumnya dan sudah dikelompokkan kedalam 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.

Sebelum menghitung nilai gain dari atribut, hitung dahulu nilai entropy. Untuk menghitung nilai entropy digunakan rumus :

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

Keterangan :

S = Himpunan Kasus

n = jumlah partisi S

P_i = Proporsi S_i terhadap S

3. Kemudian hitung nilai gain menggunakan rumus :

$$Gain(S, A) =$$

$$entropy(S) - \sum_{i=1}^n \frac{|S_i|}{S} * Entropy(S_i) \quad (2)$$

Keterangan :

S = Himpunan Kasus

A = Fitur

n = Jumlah partisi atribut A

|S_i| = Proporsi S_i terhadap S

|S| = Jumlah kasus dalam S

4. Ulangi langkah ke-2 hingga semua record terpartisi.

5. Proses partisi pohon keputusan akan berhenti saat :

a. Semua record dalam simpul N mendapat kelas yang sama.

b. Tidak ada atribut didalam record yang dipartisi lagi.

c. Tidak ada record di dalam cabang yang kosong.

METODE

Metode pengumpulan data pada penelitian ini :

1. Observasi

Observasi dilakukan dengan mendatangi langsung ke yayasan baitul maal dimana penulismelihat dan mengamati objek secara langsung

bagaimanasistem seleksi penerimaan beasiswa dilakukan.

2. Wawancara

Dalam metode ini, penulis mengumpulkan data dengan mewawancarai secara langsung pegawai yayasan baitul maal untuk mendapatkan data yang berkaitan dengan penelitian ini.

3. Studi Pustaka

Penulis mencari beberapa data yang diambil dari buku-buku, jurnal dan referensi lainnya yang menunjang penelitian ini.

Penelitian ini merupakan jenis penelitian kuantitatif, dengan menggunakan 18 variabel bebas dan 1 variabel tidak bebas.

Dalam penelitian ini menggunakan proses KDD, yang terdiri dari 5 tahap (Hermawati, 2013):

1. Selection

Data yang didapat dari baitul Maal adalah data calon beasiswa tahun 2010, jumlah data 1535 data. Atribut variabel yang ada sebanyak 18 atribut.

2. Preprocessing

Setelah dilakukan preprocessing data, atribut atau variabel yang digunakan terdiri dari 12 atribut yang terdapat dalam penentuan beasiswa. Dengan 11 atribut predictor dan 1 atribut tujuan. Atribut yang menjadi parameter beasiswa dapat dilihat sebagai berikut :

Tabel 1. Atribut dan nilai katagori

No	Atribut	Nilai
1	Usia	21-55 >55
2	Pekerjaan	Wiraswasta Pekerja tetap pekerja tidak tetap tidak bekerja
3	Pendidikan	SD SMP SMA
4	Kelas	1-6 7-9 10-12
5	Kepemilikan rumah tinggal	Rumah sewa

	Rumah sendiri
	Menumpang
6	Pendapatan
	<500000
	500000-1000000
	1050000-2000000
	>2000000
	Tidak ada
7	Pengeluaran
	<500000
	500000-1000000
	1050000-2000000
	>2000000
	Tidak tentu
8	Tanggung jawab keluarga
	1
	2-3
	>3
9	Peringkat
	1
	2-3
	4-10
	tdk ada
10	Organisasi
	Aktif Organisasi
	tdk ada
11	Nilai Raport
	<60
	60-70
	71-80
	>80
	tidak ada

3. Transformation

Dalam tahap ini, dilakukan eliminasi derau, outliers, missing value. Setelah dilakukan proses tersebut, data yang

digunakan menjadi 1120 data dan dengan 12 variabel, seperti pada gambar berikut :

Usia	Pekerjaan	Pendidikan	Kls	Kepemilikan Rumah Tinggal	Pendapatan	Pengeluaran	Tanggungans Keluarga	Peringkat	Organisasi	Nilai Raport	Remark
21-55	Wiraswasta	SD	1-6	Rumah Sewa	<500000	1050000-2000000	2-3	4-10	tdk ada	>80	yes
21-55	Wiraswasta	SD	1-6	Rumah Sewa	500000-1000000	500000-1000000	>3	tdk ada	tdk ada	71-80	yes
21-55	Wiraswasta	SD	1-6	Rumah Sewa	500000-1000000	500000-1000000	2-3	2-3	tdk ada	>80	yes
21-55	Wiraswasta	SD	1-6	Rumah Sewa	1050000-2000000	1050000-2000000	2-3	2-3	tdk ada	>80	no
21-55	Wiraswasta	SD	1-6	Rumah Sendiri	1050000-2000000	1050000-2000000	>3	4-10	tdk ada	60-70	no
21-55	Tidak Bekerja	SD	1-6	Rumah Sewa	500000-1000000	tidak tentu	2-3	tdk ada	Aktif Organisasi	60-70	yes
21-55	Wiraswasta	SD	1-6	Rumah Sendiri	500000-1000000	500000-1000000	1	2-3	tdk ada	71-80	yes
21-55	Pekerja Tidak Tetap	SMP	7-9	Rumah Sendiri	500000-1000000	500000-1000000	2-3	tdk ada	Aktif Organisasi	60-70	no
>55	Wiraswasta	SMP	7-9	Menumpang	500000-1000000	500000-1000000	2-3	1	tdk ada	71-80	yes
21-55	Pekerja Tidak Tetap	SMP	7-9	Rumah Sendiri	1050000-2000000	1050000-2000000	>3	tdk ada	Aktif Organisasi	60-70	no
21-55	Pekerja Tidak Tetap	SMA	10-12	Rumah Sendiri	1050000-2000000	1050000-2000000	>3	2-3	tdk ada	>80	yes
21-55	Pekerja Tetap	SMP	7-9	Rumah Sendiri	500000-1000000	500000-1000000	2-3	2-3	tdk ada	60-70	yes
21-55	Pekerja Tidak Tetap	SMA	10-12	Rumah Sendiri	<500000	500000-1000000	>3	tdk ada	Aktif Organisasi	71-80	yes
21-55	Pekerja Tidak Tetap	SMA	10-12	Menumpang	<500000	500000-1000000	>3	tdk ada	Aktif Organisasi	71-80	yes
21-55	Wiraswasta	SMP	7-9	Rumah Sendiri	500000-1000000	<500000	2-3	tdk ada	Aktif Organisasi	60-70	no
21-55	Wiraswasta	SD	1-6	Rumah Sewa	500000-1000000	1050000-2000000	2-3	2-3	tdk ada	71-80	yes
21-55	Pekerja Tidak Tetap	SMP	7-9	Rumah Sendiri	500000-1000000	<500000	2-3	tdk ada	tdk ada	60-70	no

Gambar 1. Data Hasil Tahap Transformation

4. Data mining

Pada tahap ini, dilakukan pemrosesan data yang diklasifikasikan oleh model dan kemudian menghasilkan sejumlah aturan. Pada penelitian ini menggunakan algoritma C4.5 menggunakan langkah-langkah dalam algoritma C4.5.

5. Interpretation/evaluation

Pada tahap ini dilakukan analisa pola yang didapat dari data mining.

HASIL DAN PEMBAHASAN

Penelitian ini bertujuan untuk menentukan kriteria pemberian beasiswa dengan menggunakan algoritma C4.5, dengan langkah-langkah sebagai berikut :

1. Siapkan data training. Data training yang digunakan dalam tahap ini berjumlah 1120 record.
2. Hitung nilai entropy.

Berdasarkan menggunakan rumus (1) pada metode penelitian untuk perhitungan entropy, didapat entropy sebagai berikut :

$$entropy(S) = (-571/1120 * \log_2(571/1120)) + (-549/1120 * \log_2(549/1120)) = 0.999721656$$

3. Setelah itu, hitung nilai gain untuk setiap atribut menggunakan rumus (2) pada metode penelitian, lalu pilih nilai gain yang tertinggi. Nilai gain yang tertinggi yang dijadikan akar dari pohon. Contoh perhitungan gain untuk atribut usia :
Gain (S,A) = 0.99-
((1029/1120*0.999)+(91/1120*0.9978))= 0.0000796

Berikut hasil perhitungan nilai entropy dan gain, sebagai berikut :

Tabel 2. Hasil Entropy dan Gain

simpul	kasus	Yes	No	entropy	gain
Usia	1120	571	549	0.9997	0.0000796
21-55	1029	523	506	0.9998	
>55	91	48	43	0.9978	
Pekerjaan	1120	571	549		0.008979
Wiraswasta	495	279	216	0.9883	
Pekerja tetap	137	57	80	0.9796	
pekerja tidak tetap	417	206	211	0.9999	
tidak bekerja	71	29	42	0.9757	
Pendidikan	1120	571	549		-0.0298363
SD	451	231	220	1.0356	
SMP	371	181	190	0.9654	
SMA	298	159	139	1.1002	
Kelas	1120	571	549		0.0009661
1-6	452	232	220	0.9995	
7-9	370	180	190	0.9995	
10-12	298	159	139	0.9967	
Kepemilikan rumah tinggal	1120	571	549		0.0016925
Rumah sewa	161	84	77	0.9986	
Rumah sendiri	675	354	321	0.9983	
Menumpang	284	133	151	0.9971	
Pendapatan	1120	571	549		0.0058606
<500000	217	127	90	0.9789	
500000-1000000	709	359	350	0.9999	
1050000-2000000	167	73	94	0.9886	
>2000000	6	3	3	1.0000	
Tidak ada	21	9	12	0.9852	
Pengeluaran	1120	571	549		0.0026098
<500000	194	97	97	1.0000	
500000-1000000	503	270	233	0.9961	
1050000-2000000	326	161	165	0.9999	
>2000000	26	13	13	1.0000	
Tidak tentu	71	30	41	0.9826	
Tanggungjawab keluarga	1120	571	549		0.0001041
1	145	74	71	0.9997	
2-3	679	349	330	0.9994	
>3	296	148	148	1.0000	
Peringkat	1120	571	549		0.0258829
1	116	82	34	0.8727	
2-3	131	81	50	0.9592	
4-10	129	76	53	0.9769	
tdk ada	744	332	412	0.9916	
Organisasi	1120	571	549		0.0192509
Aktif Organisasi	572	246	326	0.9858	
tdk ada	548	325	223	0.9749	
Nilai Raport	1120	571	549		0.1503534
<60	10	0	10	0.0000	
60-70	315	85	230	0.8412	
71-80	576	365	211	0.9478	
>80	155	118	37	0.7929	

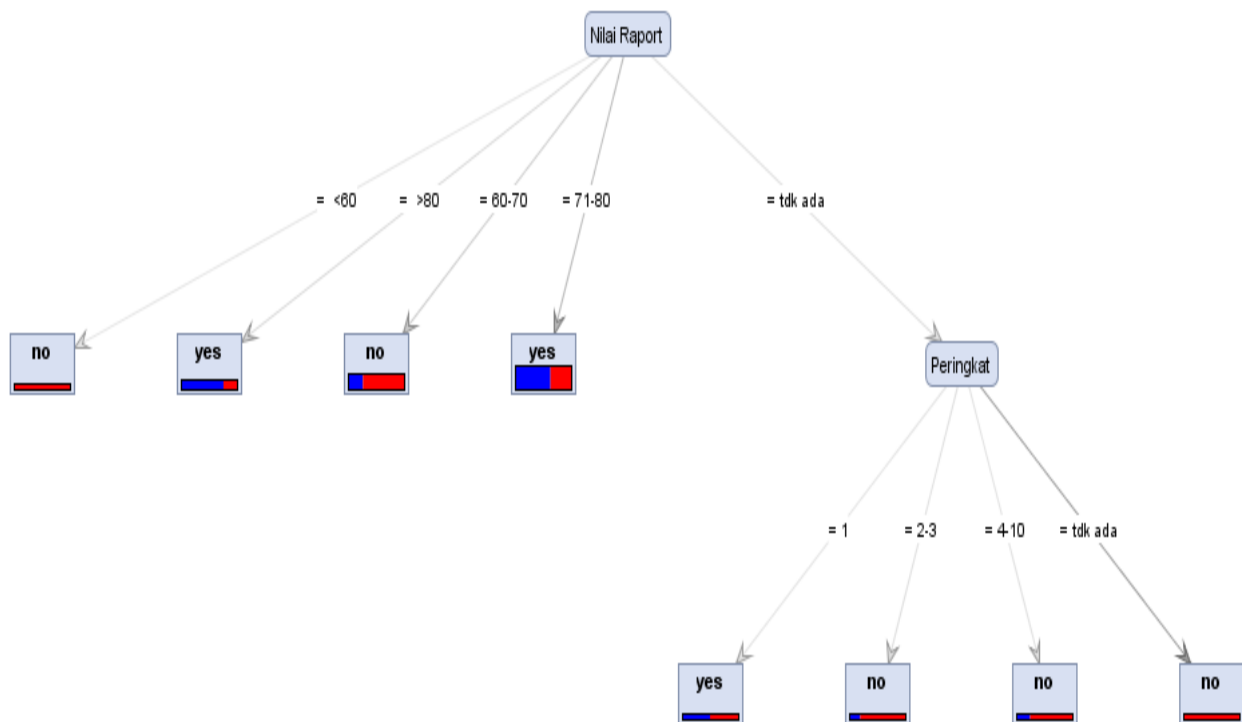
tidak ada	64	3	61	0.2730
	1120	571	549	

Dari Tabel 2, terlihat bahwa atribut nilai raport mempunyai nilai gain tertinggi yaitu 0.150353351. Oleh karena itu, nilai raport dijadikan simpul akar pada pohon keputusan.

Selanjutnya menentukan simpul berikutnya, yaitu simpul 1.1, dengan

dilakukan perhitungan entropy dan gain masing-masing atribut dari nilai raport.

Setelah dilakukan perhitungan entropy dan gain untuk seluruh atribut dalam kelas, maka akan didapat sebuah pohon keputusan seperti dibawah ini :



Gambar 2. Pohon keputusan menggunakan algoritma C4.5

Pada gambar 2, yaitu pohon keputusan yang dihasilkan dari perhitungan entropy dan gain untuk keseluruhan atribut. Berdasarkan hasil perhitungan entropy dan gain pada tabel 2, nilai raport mempunyai nilai gain tertinggi dan menghasilkan empat simpul sesuai dengan atribut yang dimilikinya, yaitu : <60, 60-70, 71-80, >80, dan tidak ada.

Kemudian hitung entropy dan gain masing-masing atribut dari nilai raport. Berdasarkan perhitungan didapat nilai gain tertinggi yaitu peringkat menjadi simpul pada level berikutnya. Pada gambar 2 terlihat atribut nilai raport yang memiliki simpul tidak ada, akan

membentuk cabang kedua, yaitu: =1, =2-3, =4-10, dan tidak ada. Berdasarkan pohon keputusan pada gambar 2, dapat membentuk aturan-aturan, yaitu sebagai berikut :

- a. R1: If Nilai Raport tidak ada AND Peringkat tidak ada THEN no
- b. R2: If Nilai Raport tidak ada AND Peringkat 4-10 THEN no
- c. R3: If Nilai Raport tidak ada AND Peringkat 2-3 THEN no
- d. R4: If Nilai Raport tidak ada AND Peringkat 1 THEN yes
- e. R5: If Nilai Raport 71-80 THEN yes
- f. R6: If Nilai Raport 60-70 THEN no
- g. R7: If Nilai Raport ≥ 80 THEN yes
- h. R8: If Nilai Raport ≤ 60 THEN no

Sehingga berdasarkan rule diatas dihasilkan pengambilan keputusan yang beasiswa berdasarkan kriteria:

1. Jika siswa atau siswi Nilai Raportnya tidak ada dan Peringkatnya tidak ada maka siswa atau siswi tersebut tidak mendapatkan beasiswa.
2. Jika siswa atau siswi Nilai Raportnya tidak ada dan Peringkatnya 4-10 maka siswa atau siswi tersebut tidak mendapatkan beasiswa.
3. Jika siswa atau siswi Nilai Raportnya tidak ada dan Peringkatnya 2-3 maka siswa atau siswi tersebut tidak mendapatkan beasiswa.
4. Jika siswa atau siswi Nilai Raport tidak ada dan Peringkatnya 1 maka siswa atau siswi tersebut mendapatkan beasiswa.

5. Jika siswa atau siswi Nilai Raport 71-80 maka siswa atau siswi tersebut mendapatkan beasiswa.
6. Jika siswa atau siswi Nilai Raport 60-70 maka siswa atau siswi tersebut tidak mendapatkan beasiswa.
7. Jika siswa atau siswi Nilai Raport ≥ 80 maka siswa atau siswi tersebut mendapatkan beasiswa.
8. Jika siswa atau siswi Nilai Raport ≤ 60 maka siswa atau siswi tersebut mendapatkan beasiswa.

Tahap akhir dalam penelitian ini dengan melakukan tahap pengujian tingkat akurasi dengan menggunakan confusion matrix dan kurva ROC/AUC (*Area Under Cover*), dengan rapidminer. Berikut hasil Confusion matrix dan kurva ROC/AUC (*Area Under Cover*), dengan rapidminer :

accuracy: 69,73% +/- 3,49% (mikro: 69,73%)			
	true yes	true no	class precision
pred. yes	483	251	65,80%
pred. no	88	298	77,20%
class recall	84,59%	54,28%	

Gambar 3. Confusion matrix algoritma C4.5

Berdasarkan gambar 3. Perhitungan akurasi data training menggunakan algoritma C4.5. Diketahui dari 1120 data dan 11 atribut (usia, pekerjaan, pendidikan, kelas, kepemilikan rumah tinggal, pendapatan, pengeluaran, tanggungan keluarga, peringkat, organisasi, nilai raport), dengan menggunakan algoritma C4.5 menghasilkan akurasi sebesar 69,73%, didapat 483 data prediksi yes sesuai dengan yes, 251 prediksi yes ternyata no, 88 data prediksi no ternyata yes dan 298 data predisi no sesuai dengan no.

SIMPULAN DAN SARAN

Penelitian ini digunakan untuk mendukung keputusan pemberian beasiswa dengan menggunakan algoritma C4.5 yang dapat membantu pemberian beasiswa bagi siswa atau siswi sekolah

dasar (SD), Sekolah Lanjutan Tingkat Atas (SLTP), dan Sekolah Lanjutan Tingkat Atas (SLTA). Dalam penelitian ini didapat kriteria penentu utama beasiswa yaitu nilai raport, jika nilai raportnya tidak ada maka dengan melihat data peringkat, jika peringkat 1 maka dapat beasiswa, tetapi jika tidak ada peringkat tidak mendapat beasiswa. Atribut yang lainnya tidak termasuk dalam penentuan beasiswa. Penelitian ini melakukan pengujian dengan menggunakan confusion matrix dan kurva ROC/AUC (*Area Under Cover*), dengan hasil 69,73%.

Saran dari penelitian ini :

1. Penelitian ini dilakukan dengan menggunakan data calon beasiswa dari satu lembaga saja. Apabila dilakukan penelitian yang sama dapat dilakukan di lembaga atau perusahaan lain dengan atribut yang bervariasi,

sehingga mendapatkan hasil yang berbeda.

2. Menambahkan jumlah atribut yang lebih banyak, sehingga diperoleh pengukuran yang lebih baik lagi.
3. Menggunakan model data mining yang lainnya, agar didapat hasil yang berbeda dengan tingkat akurasi yang berbeda.

DAFTAR RUJUKAN

Ahda, Fadhli Almu'iini dan Saiful Bahri. 2017. Sistem Pendukung Keputusan Pemberian Beasiswa Untuk Siswa Kurang Mampu Di SMK Muhammadiyah 1 Kepanjen Menggunakan Metode Fuzzy Tsukamoto. *Jurnal Spirit*. Vol 9 No.2 Nopember 2017, hal 48-56. ISSN : 2085-3092. STMIK ASIA Malang.

Arikunto, Suharsimi. 2014. *Prosedur Penelitian*. Jakarta: Rineka Cipta.

Guterres, Jose Augusto Duarte. 2015. Kelayakan Algoritma C4.5 Sebagai Pendukung Keputusan Dalam Pengajuan Penerimaan Beasiswa. *Seminar Nasional Teknologi Informasi*.

Haryati, Siska, Aji Sudarsono, Eko Suryana. 2015. Implementasi Data Mining Untuk Memprediksi Masa Studi Mahasiswa Menggunakan Algoritma C4.5 (Studi Kasus Universitas Dehasen Bengkulu). *Jurnal Media Infotama*, Universitas Dehasen Bengkulu.

Hermawati, Fajar Astuti. 2013. *Data Mining*. Yogyakarta: Andi Offset.

Hijriana, Nadiya dan Muhammad Rasyidan. 2017. Penerapan Metode Decision Tree Algoritma C4.5 Untuk Seleksi Calon Penerima Beasiswa Tingkat Universitas. *AI Ulum Sains dan Teknologi* Vol. 3 No. 1 Nopember 2017.

Muslihudin, Muhamad. 2017. *Sistem Pendukung Keputusan Menentukan*

Kelayakan Penerimaan Bantuan Pengusaha Ayam Petelur Oleh Dinas Peternakan Kabupaten Pesawaran Menggunakan Metode Simple Additive Weighting. *Jurnal Pendidikan Teknologi dan Kejuruan* Vol.14, No.2, Juli 2017, Hal:120. P-ISSN : 0216-3241 ; E-ISSN : 2541-065.

Nurtanto, M, M, Fawaid, Y. Fargianti, S.D. Ramdani, dan S.Nurhaji. Faktor Pengaruh Minat Masuk Perguruan Tinggi Di SMK Serang. *Jurnal Pendidikan Teknologi dan Kejuruan* Vol. 14, No. 1, Januari 2017, Hal : 11 P-ISSN : 0216-3241 ; E-ISSN : 2541-0652.

Praja, Mahindra Suryaning dan Erna Zuni Astuti. 2016. Penerapan Data Mining Untuk Rekomendasi Beasiswa Pada SMA N 1 Mlonggo Menggunakan Algoritma C4.5. Universitas Dian Nuswantoro, Semarang, skripsi 2016.

Prasetyo, Eko. 2012. *Data Mining Konsep dan Aplikasi Menggunakan MATLAB*. Yogyakarta: Andi Yogyakarta.

Rainer, dedi. 2017. Pengertian Beasiswa, Tujuan, Syarat, Manfaat, Jenis, Contoh Terlengkap. <http://www.studineews.co.id>.

Suryadi, Achmad. 2007. *Analisa Manajemen Pengambilan Keputusan*. Surabaya: Graha Ilmu.

Vulandari, Retno Tri. 2017. *Data Mining Teori dan Aplikasi Rapidminer*. Yogyakarta: Gava Media.