



Data Mining Untuk Seleksi Kelayakan Calon Penerima Beasiswa di STMIK Bina Bangsa Kendari

Nur Islamuddin

STMIK Bina Bangsa Kendari
islanur@gmail.com

Abstrak

Program beasiswa yang diperuntukkan bagi siswa SMA/SMK/Madrasah Aliyah telah ada di STMIK Bina Bangsa Kendari sejak tahun 2010. Beberapa tahapan dan kriteria dalam penentuan beasiswa terkadang mengakibatkan sulitnya membuat keputusan dalam menentukan calon penerima beasiswa. Data calon penerima disajikan dalam bentuk tabel akan tetapi tidak mudah dalam proses membaca tabel dengan cepat dan tepat ditambah lagi dengan jumlah data yang cukup besar. Tujuan dari penelitian mengumpulkan informasi tentang data calon penerima beasiswa penuh dalam satu tabel mulai dari data-data penanggung calon penerima beasiswa hingga hasil nilai UN calon penerima beasiswa. Informasi disajikan dengan cara mengklasifikasikan calon penerima beasiswa penuh menggunakan algoritma C4.5 dengan mencari nilai Entropy dan Gain dalam membentuk pohon keputusan. Dengan menggunakan sistem informasi pembuatan pohon keputusan dari data calon mahasiswa menjadi lebih mudah dan cepat sehingga dapat melihat factor apa saja yang menjadi pertimbangan dalam pengambilan keputusan.

Kata Kunci: Data mining, Beasiswa penuh, Algoritma C4.5, Model pendukung keputusan.

Abstract

The scholarship program that is intended for high school / vocational / Madrasah Aliyah students has been in the STMIK Bina Bangsa Kendari since 2010. Several stages and criteria in determining scholarships sometimes make it difficult to make decisions in determining prospective scholarship recipients. Prospective recipient data is presented in tabular form but it is not easy in the process of reading tables quickly and precisely coupled with a sizable amount of data. The purpose of this research is to collect information about the data of prospective scholarship recipients in one table starting from the data of the guarantor of prospective scholarship recipients to the results of UN scores of prospective scholarship recipients. Information is presented by classifying prospective full scholarship recipients using the C4.5 algorithm by finding the value of Entropy and Gain in forming a decision tree. By using a decision tree information system making data from prospective students becomes easier and faster so that you can see what factors are considered in making decisions.

Keywords: Data mining, Full scholarships, C4.5 Algorithm, Decision support models.

1. Pendahuluan

STMIK Bina Bangsa Kendari merupakan sebuah Perguruan Tinggi Swasta, yang berdiri sejak tahun 2004. Pada tahun 2010 STMIK Bina Bangsa memulai membuka jalur pendaftaran beasiswa penuh bagi masyarakat umum untuk lulusan SMU/SMK/MA dan sederajat.

Beasiswa merupakan salah satu solusi untuk mengatasi masalah biaya bagi mereka yang kurang mampu. Menurut Murniasih dalam [1] 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: mereka tidak perlu lagi khawatir dengan tagihan biaya pendidikan dan anak tersebut berusaha untuk mendapatkan nilai dan prestasi yang tinggi.

Salah satu permasalahan dalam penyaluran dana beasiswa adalah masih adanya pemberian beasiswa yang kurang tepat sasaran, sehingga menyebabkan tidak maksimalnya pendayagunaan dana beasiswa tersebut. Selain itu proses seleksi calon penerima beasiswa yang dilakukan secara manual terkadang membutuhkan waktu cukup lama sehingga kurang efisien.

Beberapa penelitian telah dilakukan untuk membantu pengambilan keputusan dalam proses seleksi calon penerima beasiswa, dengan harapan agar beasiswa dapat diberikan kepada mahasiswa yang memang pantas mendapatkannya atau tepat sasaran, serta dapat memberikan efisiensi waktu dalam proses pengambilan keputusan. Salah satu metode yang sering digunakan adalah metode Decision Tree Algoritma C4.5 seperti penelitian yang dilakukan oleh Muh. Safri Juliardi, dkk. dengan kesimpulan bahwa Algoritma C4.5 dapat digunakan sebagai metode alternatif untuk membantu proses seleksi penerima beasiswa Bidikmisi[2].

Selain itu, Jose Augusto Duarte Guterres juga melakukan perbandingan metode untuk klasifikasi calon penerima beasiswa. Hasilnya adalah bahwa Algoritma C4.5 layak dijadikan sebagai pendukung keputusan dalam proses pengajuan penerima beasiswa pada Stikom Artah Buana Kupang karena hasil persentasi dari Algoritma C4.5 yang lebih tinggi daripada metode klasifikasi yang lainnya[3].

Data Mining merupakan teknologi baru yang sangat berguna untuk menemukan informasi yang sangat penting dari gudang data. Ketersediaan data yang banyak dan kebutuhan akan informasi atau pengetahuan sebagai pendukung pengambilan keputusan untuk membuat solusi bisnis dan dukungan infrastruktur di bidang sistem informasi merupakan cikal-bakal dari lahirnya teknologi data mining. Sehingga informasi tersebut bisa digunakan sebagai solusi pengambilan keputusan di dunia bisnis, untuk pengembangan bisnis. Konsep teknologi informasi yang berkaitan dengan data dan informasi adalah bagian dari data mining. Konsep data mining adalah sebagai proses atau teknik pemodelan yang mempergunakan analisis dengan variasi data yang banyak untuk mendapatkan pola dan hubungan diantara variasi data tersebut. Kehadiran data mining dilatarbelakangi dengan adanya masalah data explosion atau ledakan data yang dialami oleh banyak organisasi yang telah mengumpulkan data sekian tahun lamanya.

Dari latar belakang permasalahan tersebut peneliti mencoba mengembangkan suatu sistem informasi untuk penyajian data dari hasil lolos seleksi wawancara yang berisi faktor-faktor apa saja dari rekomendasi pewawancara sebagai pendukung keputusan. Penelitian ini menerapkan algoritma C4.5 dalam sebuah sistem informasi yang dapat mengetahui pola pengambilan keputusan penerimaan beasiswa. Dari pola yang dihasilkan akan didapatkan sebuah informasi atau knowledge.

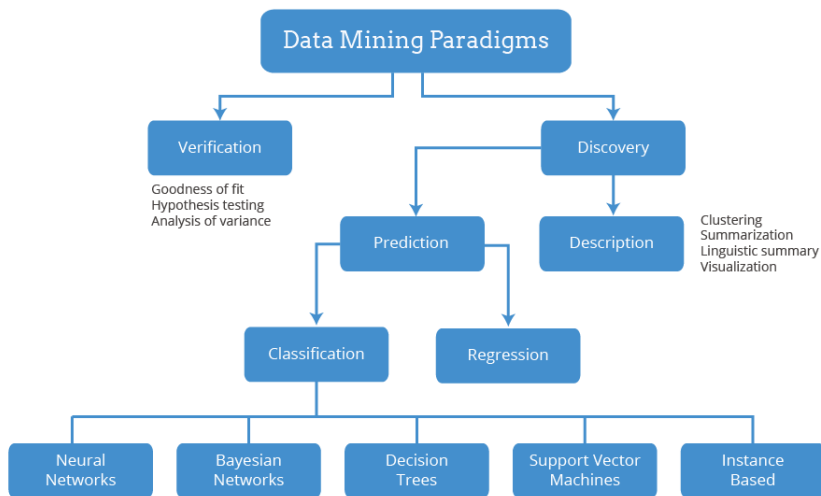
Dengan mengacu pada latar belakang masalah diatas, maka peneliti akan membahas permasalahan tentang “Bagaimana memanfaatkan algoritma C4.5 Data Mining untuk model penentuan pengambilan keputusan calon penerima beasiswa penuh di STMIK Bina Bangsa”.

Kegiatan penelitian ini bertujuan untuk menerapkan teknik data mining dalam menampilkan informasi untuk pengambilan keputusan mengenai calon penerima beasiswa penuh di STMIK Bina Bangsa. Dan dengan penerapan data mining ini diharapkan dapat memberikan pengetahuan yang sebelumnya tersembunyi didalam data sehingga menjadi informasi berharga, berupa informasi yang dapat membantu memprediksi dan mengambil keputusan untuk calon penerima beasiswa penuh di STMIK Bina Bangsa Kendari.

2. Metode

2.1. Klasifikasi

Klasifikasi atau taksonomi adalah proses menempatkan suatu objek atau konsep kedalam satu set kategori berdasarkan objek atau konsep yang bersangkutan. Ada banyak metode data mining yang digunakan untuk tujuan yang berbeda-beda. Metode klasifikasi digunakan untuk membantu dalam memahami pengelompokan data. Klasifikasi sendiri merupakan cabang dari discovery data mining seperti yang ditunjukkan pada gambar 1.



Gambar 1. Data Mining Taksonomi

2.2. Algoritma C4.5

Algoritma C4.5 dan pohon keputusan merupakan dua model yang tak terpisahkan, karena untuk membangun sebuah pohon keputusan, dibutuhkan algoritma C4.5. Di akhir tahun 1970 hingga di awal tahun 1980-an, J. Ross Quinlan seorang peneliti di bidang mesin pembelajaran mengembangkan sebuah model pohon keputusan yang dinamakan ID3 (Iterative Dichotomiser), walaupun sebenarnya proyek ini telah dibuat sebelumnya oleh E.B. Hunt, J. Marin, dan P.T. Stone. Kemudian Quinlan membuat algoritma dari pengembangan ID3 yang dinamakan C4.5 yang berbasis supervised learning.

Menurut (Witten, Frank, & Hall, 2011) serangkaian perbaikan yang dilakukan pada ID3 mencapai puncaknya dengan menghasilkan sebuah sistem praktis dan berpengaruh untuk decision tree yaitu C4.5. Perbaikan ini meliputi metode untuk menangani numeric attributes, missing values, noisy data, dan aturan yang menghasilkan rules dari trees.

Ada beberapa tahapan dalam membuat sebuah pohon keputusan dalam algoritma C4.5 (Larose, 2005) yaitu :

1. Mempersiapkan data training. Data training biasanya diambil dari data histori yang pernah terjadi sebelumnya atau disebut data masa lalu dan sudah dikelompokkan dalam kelas-kelas tertentu.
2. Menghitung akar dari pohon. Akar akan diambil dari atribut yang akan 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 :

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

Keterangan :

S = Himpunan (dataset) kasus

n = Jumlah partisi S

Pi = Proporsi S-i terhadap S

Kemudian hitung nilai gain menggunakan rumus:

$$Gain(S, A) =$$

$$Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{S} \times Entropy(S_i)$$

Keterangan :

S = Himpunan (dataset) kasus

A = Fitur

n = Jumlah partisi atribut A

|S_i| = Proporsi S_i terhadap S

|S| = Jumlah kasus dalam S

3. Ulangi langkah 2 dan langkah ke 3 hingga semua record terpartisi
4. 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 didalam cabang yang kosong

Metode pengumpulan data pada penelitian ini:

1. Studi pustaka

Yaitu mencari informasi dari beberapa buku referensi, serta situs di internet yang dapat membantu memperoleh pengetahuan dan pokok permasalahan yang berhubungan dengan materi penulisan ini.

2. Pengumpulan data SPMB

Mengumpulkan data penerima hasil seleksi wawancara yang diambil dari database Sistem Informasi Penerimaan Mahasiswa Baru, data yang diambil untuk dilakukan proses mining.

3. Metode analisis data

Adapun untuk menganalisis data dalam penerapan data mining ini menggunakan tahapan Knowledge Discovery in Databases (KDD).

Proses KDD, yang terdiri dari 5 tahap (Hermawati, 2013): yang terdiri dari beberapa tahapan, yaitu data selection, preprocessing, transformation, data mining, dan evaluation.

1. Data Selection

Data yang didapat dari pangkalan data STMIK Bina Bangsa Kendari adalah data calon penerima beasiswa tahun 2016, tahun 2017 dan tahun 2018, dengan jumlah 95 data. Atribut variabel yang ada sebanyak 18 atribut.

2. Data Praproses

Setelah dilakukan preprocessing data, atribut atau variabel yang digunakan terdiri dari 9 atribut yang terdapat dalam penentuan beasiswa. Dengan 8 atribut predictor dan 1 atribut tujuan dan beberapa pengelompokan data sebagai berikut:

21-55	Muda
>55	Tua

Tabel 1. Usia Penanggung

<500.000	Rendah
500.000-1.000.000	Cukup
1.050.000-2.000.000	Rata-rata
>2.000.000	Tinggi

Tabel 2. Penghasilan Penanggung

1	Sedikit
2-3	Sedang
>3	Banyak

Tabel 3. Jumlah Tanggungan

<60	Kurang
60-70	Cukup
71-80	Baik
>80	Sangat Baik

Tabel 4. Nilai UN

Dari pengelompokan data yang telah dilakukan, maka atribut yang menjadi parameter beasiswa dapat dilihat sebagai berikut :

No	Atribut	Nilai
1	Penanggung	ORANG TUA
		WALI
		DIRI SENDIRI
2	Jenis Kelamin Penanggung	LAKI-LAKI
		PEREMPUAN
3	Usia Penanggung	MUDA
		TUA
4	Pekerjaan Penanggung	PEKERJA TETAP
		PEKERJA TIDAK TETAP
		TIDAK BEKERJA
5	Kepemilikan Rumah Tinggal	RUMAH SEWA
		RUMAH SENDIRI
		MENUMPANG
6	Penghasilan Penanggung	RENDAH
		CUKUP
		RATA-RATA
		TINGGI
7	Jumlah Tanggungan Penanggung	SEDIKIT
		SEDANG
		BANYAK
8	Nilai UN	KURANG
		CUKUP
		BAIK
		SANGAT BAIK

Tabel 5. Atribut dan Nilai Kategori

3. Transformation

Dalam tahap ini, dilakukan eliminasi derau, outliers, missing value. Setelah dilakukan proses tersebut, data yang digunakan menjadi 95 data dan dengan 9 variabel, seperti pada gambar berikut :

NO DATA	PENANGGUNG	JENIS KELAMIN	USIA	PEKERJAAN	KEPEMILIKAN RUMAH	PENGHASILAN	JUMLAH TANGGUNGAN	NILAI UN	TERIMA
DATA KE-1	ORANG TUA	LAKI-LAKI	MUDA	PEKERJA TIDAK TETAP	RUMAH SENDIRI	RATA-RATA	SEDANG	KURANG	YA
DATA KE-2	ORANG TUA	LAKI-LAKI	MUDA	PEKERJA TETAP	RUMAH SEWA	CUKUP	SEDANG	BAIK	YA
DATA KE-3	ORANG TUA	LAKI-LAKI	MUDA	PEKERJA TETAP	RUMAH SENDIRI	CUKUP	SEDANG	SANGAT BAIK	TIDAK
DATA KE-4	ORANG TUA	LAKI-LAKI	TUA	PEKERJA TETAP	RUMAH SENDIRI	TINGGI	BANYAK	KURANG	TIDAK
DATA KE-5	WALI	PEREMPUAN	MUDA	PEKERJA TETAP	RUMAH SENDIRI	CUKUP	BANYAK	SANGAT BAIK	YA
DATA KE-6	WALI	LAKI-LAKI	MUDA	TIDAK BEKERJA	RUMAH SEWA	RENDAH	SEDIKIT	CUKUP	YA
DATA KE-7	ORANG TUA	LAKI-LAKI	MUDA	PEKERJA TIDAK TETAP	RUMAH SENDIRI	RENDAH	SEDANG	KURANG	YA
DATA KE-8	ORANG TUA	PEREMPUAN	MUDA	PEKERJA TETAP	RUMAH SEWA	RATA-RATA	SEDANG	CUKUP	TIDAK
DATA KE-9	WALI	LAKI-LAKI	MUDA	PEKERJA TETAP	RUMAH SEWA	RATA-RATA	SEDIKIT	KURANG	TIDAK
DATA KE-10	ORANG TUA	LAKI-LAKI	MUDA	PEKERJA TIDAK TETAP	RUMAH SEWA	CUKUP	SEDANG	KURANG	YA
DATA KE-11	WALI	LAKI-LAKI	TUA	PEKERJA TIDAK TETAP	RUMAH SENDIRI	CUKUP	BANYAK	BAIK	YA
DATA KE-12	ORANG TUA	PEREMPUAN	MUDA	PEKERJA TETAP	RUMAH SENDIRI	RATA-RATA	SEDANG	BAIK	TIDAK
DATA KE-13	ORANG TUA	LAKI-LAKI	TUA	PEKERJA TIDAK TETAP	RUMAH SENDIRI	RATA-RATA	SEDANG	KURANG	YA
DATA KE-14	WALI	LAKI-LAKI	MUDA	PEKERJA TETAP	RUMAH SEWA	CUKUP	SEDIKIT	KURANG	TIDAK
DATA KE-15	WALI	PEREMPUAN	MUDA	PEKERJA TETAP	RUMAH SENDIRI	CUKUP	SEDANG	BAIK	YA
DATA KE-16	ORANG TUA	LAKI-LAKI	MUDA	PEKERJA TIDAK TETAP	RUMAH SENDIRI	RENDAH	BANYAK	CUKUP	YA
DATA KE-17	ORANG TUA	PEREMPUAN	TUA	PEKERJA TIDAK TETAP	RUMAH SEWA	RENDAH	SEDANG	BAIK	YA
DATA KE-18	ORANG TUA	LAKI-LAKI	MUDA	PEKERJA TETAP	RUMAH SENDIRI	RATA-RATA	SEDIKIT	CUKUP	TIDAK
DATA KE-19	WALI	PEREMPUAN	MUDA	PEKERJA TETAP	MENUMPANG	CUKUP	BANYAK	CUKUP	YA
DATA KE-20	ORANG TUA	LAKI-LAKI	MUDA	PEKERJA TETAP	RUMAH SENDIRI	RATA-RATA	BANYAK	BAIK	TIDAK

Tabel 6. Data Transformasi

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. Evaluation

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

3. 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 95 record.

2. Hitung nilai entropy.

Dengan berdasarkan rumus (1) pada metode penelitian untuk perhitungan entropy, didapat entropy sebagai berikut:

$$\begin{aligned} \text{Entropy (S)} &= (-50/95 * \log_2(50/95)) + (-45/95 * \log_2(45/95)) \\ &= 0.9980008838723 \end{aligned}$$

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 penanggung :

$$\begin{aligned} \text{Gain (S,A)} &= 0.998001 - (((64/95) * 0.965202) + ((25/95) * 0.79504) + ((6/95) * 0)) \\ &= 0.138539 \end{aligned}$$

Berikut hasil perhitungan nilai entropy dan gain:

ATRIBUT	VALUE	JUMLAH KASUS	YA	TIDAK	ENTROPY	GAIN
TOTAL		95	50	45	0.9980009	
Penanggung						0.1385386
ORANG TUA		64	25	39	0.9652017	
WALI		25	19	6	0.7950403	
DIRI SENDIRI		6	6	0	0	
Jenis Kelamin Penanggung						0.0161576
LAKI-LAKI		60	35	25	0.9798688	
PEREMPUAN		35	15	20	0.9852281	
Usia Penanggung						0.0166211
MUDA		67	32	35	0.9985533	
TUA		28	18	10	0.940286	
Pekerjaan Penanggung						0.109041
PEKERJA TETAP		74	32	42	0.9867867	
PEKERJA TIDAK TETAP		17	14	3	0.6722948	
TIDAK BEKERJA		4	4	0	0	
Kepemilikan Rumah						0.1254985
RUMAH SEWA		24	19	5	0.7382849	
RUMAH SENDIRI		67	27	40	0.97267	
MENUMPANG		4	4	0	0	
Penghasilan Penanggung						0.5271406
RENDAH		20	20	0	0	
CUKUP		33	25	8	0.7990485	
RATA-RATA		26	5	21	0.7062741	
TINGGI		16	0	16	0	
Tanggung Penanggung						0.0076656
SEDIKIT		19	9	10	0.9980009	
SEDANG		55	28	27	0.9997615	
BANYAK		21	13	8	0.9587119	
Nilai UN						0.042102
KURANG		19	8	11	0.9819408	
CUKUP		24	9	15	0.954434	
BAIK		41	26	15	0.9474351	
SANGAT BAIK		11	7	4	0.9456603	

Tabel 6. Hasil Entropy dan Gain

Dari Tabel 6, terlihat bahwa atribut penghasilan penanggung mempunyai nilai gain tertinggi yaitu 0.5271406. Oleh karena itu, atribut penghasilan penanggung dijadikan simpul akar pada pohon keputusan. Selanjutnya menentukan simpul berikutnya, yaitu simpul 1.1, dengan dilakukan perhitungan entropy dan gain masing-masing atribut dari atribut penghasilan penanggung yaitu perhitungan untuk kategori penghasilan penanggung= ”cukup” dan penghasilan penanggung=”rata-rata”. Untuk kategori penghasilan penanggung=”rendah”, dijadikan daun dengan nilai “Ya” karena tidak ada nilai lagi pada hasil “Tidak” untuk kategori “rendah”. Demikian juga pada kategori penghasilan penanggung=“Tinggi” menjadi daun untuk nilai “Tidak”.

Hasil perhitungan entropy dan gain untuk simpul 1.1 atribut penghasilan penanggung=”Cukup” dan atribut penghasilan penanggung=”Rata-rata” dapat dilihat pada gambar 2 dan gambar 3 berikut :

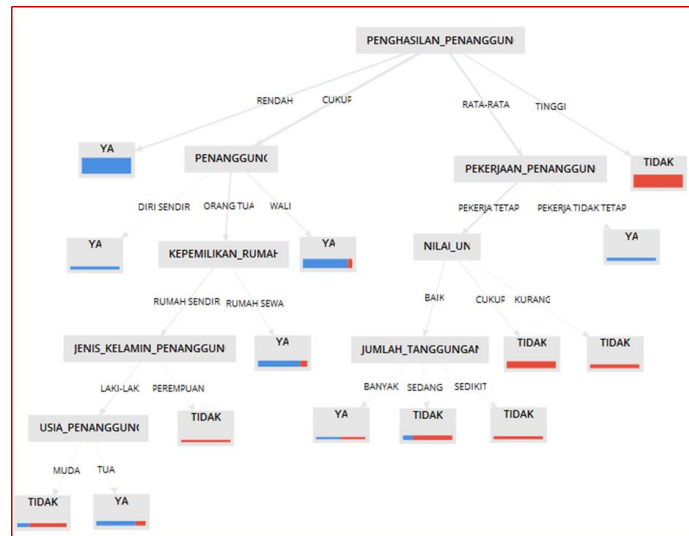
ATRIBUT	VALUE	Jumlah Kasus	YA	TIDAK	Entropy	Gain
PENGHASILAN = CUKUP		33	25	8	0.799049	
Penanggung						0.122709
	ORANG TUA	18	11	7	0.964079	
	WALI	12	11	1	0.413817	
	DIRI SENDIRI	3	3	0	0	
Jenis Kelamin Penanggung						0.003153
	LAKI-LAKI	23	17	6	0.828056	
	PEREMPUAN	10	8	2	0.721928	
Usia Penanggung						0.020934
	MUDA	20	14	6	0.881291	
	TUA	13	11	2	0.619382	
Pekerjaan Penanggung						0.000609
	PEKERJA TETAP	24	18	6	0.811278	
	PEKERJA TIDAK TETAP	9	7	2	0.764205	
	TIDAK BEKERJA	0	0	0	0	
Kepemilikan Rumah						0.054161
	RUMAH SEWA	13	11	2	0.619382	
	RUMAH SENDIRI	18	12	6	0.918296	
	MENUMPANG	2	2	0	0	
Tanggungan Penanggung						0.011638
	SEDIKIT	4	3	1	0.811278	
	SEDANG	22	16	6	0.845351	
	BANYAK	7	6	1	0.591673	
Nilai UN						0.100089
	KURANG	8	4	4	1	
	CUKUP	5	4	1	0.721928	
	BAIK	12	11	1	0.413817	
	SANGAT BAIK	8	6	2	0.811278	

Gambar 2. Hasil Entropy dan Gain Atribut Penghasilan="Cukup"

ATRIBUT	VALUE	Jumlah Kasus	YA	TIDAK	Entropy	Gain
PENGHASILAN = RATA-RATA		26	5	21	0.706274	
Penanggung						0.116609
	ORANG TUA	23	4	19	0.666578	
	WALI	2	0	2	0	
	DIRI SENDIRI	1	1	0	0	
Jenis Kelamin Penanggung						0.050524
	LAKI-LAKI	14	4	10	0.863121	
	PEREMPUAN	12	1	11	0.413817	
Usia Penanggung						0.00094
	MUDA	20	4	16	0.721928	
	TUA	6	1	5	0.650022	
Pekerjaan Penanggung						0.329226
	PEKERJA TETAP	23	2	21	0.426229	
	PEKERJA TIDAK TETAP	3	3	0	0	
	TIDAK BEKERJA	0	0	0	0	
Kepemilikan Rumah						0.038057
	RUMAH SEWA	3	0	3	0	
	RUMAH SENDIRI	23	5	18	0.755375	
	MENUMPANG	0	0	0	0	
Tanggungan Penanggung						0.118539
	SEDIKIT	7	1	6	0.591673	
	SEDANG	16	3	13	0.696212	
	BANYAK	3	1	2	0	
Nilai UN						0.136711
	KURANG	6	2	4	0.918296	
	CUKUP	9	0	9	0	
	BAIK	11	3	8	0.845351	
	SANGAT BAIK	0	0	0	0	

Gambar 3. Hasil Entropy dan Gain Atribut Penghasilan="Rata-rata"

Setelah dilakukan perhitungan entropy dan gain untuk seluruh atribut, maka akan didapat sebuah pohon keputusan seperti gambar berikut:



Gambar 4. Pohon Keputusan menggunakan Algoritma C4.5

Pada gambar 4, yaitu pohon keputusan yang dihasilkan dari perhitungan entropy dan gain untuk keseluruhan atribut. Berdasarkan hasil perhitungan entropy dan gain pada tabel 6, atribut penghasilan penanggung mempunyai nilai gain tertinggi dan menghasilkan empat simpul sesuai dengan atribut yang dimilikinya, yaitu : rendah, cukup, rata-rata, dan tinggi. Kemudian hitung entropy dan gain masing-masing atribut dari nilai report. Berdasarkan perhitungan didapat nilai gain tertinggi yaitu peringkat menjadi simpul pada level berikutnya.

Pada gambar 4 terlihat atribut penghasilan penanggung yang memiliki simpul penanggung, akan membentuk cabang kedua, yaitu: =Diri Sendiri, =Orang Tua, dan =Wali. Simpul pekerjaan penanggung juga membentuk cabang kedua, yaitu: =Pekerja Tetap dan =Pekerja Tidak Tetap.

Berdasarkan pohon keputusan pada gambar 4, dapat membentuk aturan-aturan seperti terlihat pada gambar berikut:

Tree

```

PENGHASILAN_PENANGGUNG = CUKUP
| PENANGGUNG = DIRI SENDIRI: YA {YA=3, TIDAK=0}
| PENANGGUNG = ORANG TUA
| | KEPEMILIKAN_RUMAH = RUMAH SENDIRI
| | | JENIS_KELAMIN_PENANGGUNG = LAKI-LAKI
| | | | USIA_PENANGGUNG = MUDA: TIDAK {YA=1, TIDAK=3}
| | | | USIA_PENANGGUNG = TUA: YA {YA=4, TIDAK=1}
| | | | JENIS_KELAMIN_PENANGGUNG = PEREMPUAN: TIDAK {YA=0, TIDAK=2}
| | | KEPEMILIKAN_RUMAH = RUMAH SEWA: YA {YA=6, TIDAK=1}
| | PENANGGUNG = WALI: YA {YA=10, TIDAK=1}
PENGHASILAN_PENANGGUNG = RATA-RATA
| PEKERJAAN_PENANGGUNG = PEKERJA TETAP
| | NILAI_UN = BAIK
| | | JUMLAH_TANGGUNGAN = BANYAK: YA {YA=1, TIDAK=1}
| | | JUMLAH_TANGGUNGAN = SEDANG: TIDAK {YA=1, TIDAK=4}
| | | JUMLAH_TANGGUNGAN = SEDIKIT: TIDAK {YA=0, TIDAK=3}
| | NILAI_UN = CUKUP: TIDAK {YA=0, TIDAK=3}
| | NILAI_UN = KURANG: TIDAK {YA=0, TIDAK=4}
| PEKERJAAN_PENANGGUNG = PEKERJA TIDAK TETAP: YA {YA=3, TIDAK=0}
PENGHASILAN_PENANGGUNG = RENDAH: YA {YA=20, TIDAK=0}
PENGHASILAN_PENANGGUNG = TINGGI: TIDAK {YA=0, TIDAK=16}
    
```

Gambar 5. Aturan berdasarkan pohon keputusan

Pada gambar 5, dapat di terjemahkan kedalam rumus baru sebagai berikut:

- a. R1: IF Penghasilan Penanggung = "Cukup" AND Penanggung = "Orang Tua" AND Kepemilikan Rumah = "Rumah Sendiri" AND Jenis Kelamin Penanggung = "Laki-laki" AND Usia Penanggung = "Tua" THEN "Ya"
- b. R2: IF Penghasilan Penanggung = "Cukup" AND Penanggung = "Orang Tua" AND Kepemilikan Rumah = "Rumah Sendiri" AND Jenis Kelamin Penanggung = "Laki-laki" AND Usia Penanggung = "Muda" THEN "Tidak"
- c. R3: IF Penghasilan Penanggung = "Cukup" AND Penanggung = "Orang Tua" AND Kepemilikan Rumah = "Rumah Sendiri" AND Jenis Kelamin Penanggung = "Perempuan" THEN "Tidak"
- d. R4: IF Penghasilan Penanggung = "Cukup" AND Penanggung = "Orang Tua" AND Kepemilikan Rumah = "Rumah Sewa" THEN "Ya"
- e. R5: IF Penghasilan Penanggung = "Cukup" AND Penanggung = "Diri Sendiri" THEN "Ya"
- f. R6: IF Penghasilan Penanggung = "Cukup" AND Penanggung = "Wali" THEN "Ya"
- g. R7: IF Penghasilan Penanggung = "Rata-rata" AND Pekerjaan Penanggung = "Pekerja Tetap" AND Nilai UN = "Baik" AND Jumlah Tanggungan = "Banyak" THEN "Ya"
- h. R8: IF Penghasilan Penanggung = "Rata-rata" AND Pekerjaan Penanggung = "Pekerja Tetap" AND Nilai UN = "Baik" AND Jumlah Tanggungan = "Sedang" THEN "Tidak"
- i. R9: IF Penghasilan Penanggung = "Rata-rata" AND Pekerjaan Penanggung = "Pekerja Tetap" AND Nilai UN = "Baik" AND Jumlah Tanggungan = "Sedikit" THEN "Tidak"
- j. R10: IF Penghasilan Penanggung = "Rata-rata" AND Pekerjaan Penanggung = "Pekerja Tetap" AND Nilai UN = "Cukup" THEN "Tidak"
- k. R11: IF Penghasilan Penanggung = "Rata-rata" AND Pekerjaan Penanggung = "Pekerja Tetap" AND Nilai UN = "Kurang" THEN "Tidak"
- l. R12: IF Penghasilan Penanggung = "Rata-rata" AND Pekerjaan Penanggung = "Pekerja Tidak Tetap" THEN "Ya"
- m. R13: IF Penghasilan Penanggung = "Tinggi" THEN "Tidak"
- n. R14: IF Penghasilan Penanggung = "Rendah" THEN "Ya"

Sehingga berdasarkan aturan-aturan di atas dihasilkan pengambilan keputusan untuk yang layak menerima beasiswa penuh berdasarkan kriteria:

1. Jika calon penerima beasiswa Penghasilan Penanggungnya cukup dan penanggungnya orang tua dan kepemilikan rumah adalah rumah sendiri dan jenis kelamin penanggungnya laki-laki dan usia penanggungnya tua, maka calon penerima beasiswa tersebut layak diberikan beasiswa.
2. Jika calon penerima beasiswa Penghasilan Penanggungnya cukup dan penanggungnya orang tua dan kepemilikan rumah adalah rumah sendiri dan jenis kelamin penanggungnya laki-laki dan usia penanggungnya muda, maka calon penerima beasiswa tersebut belum layak diberikan beasiswa.
3. Jika calon penerima beasiswa Penghasilan Penanggungnya cukup dan penanggungnya orang tua dan kepemilikan rumah adalah rumah sendiri dan jenis kelamin penanggungnya perempuan, maka calon penerima beasiswa tersebut belum layak diberikan beasiswa.
4. Jika calon penerima beasiswa Penghasilan Penanggungnya cukup dan penanggungnya orang tua dan kepemilikan rumah adalah rumah sewa, maka calon penerima beasiswa tersebut layak diberikan beasiswa.
5. Jika calon penerima beasiswa Penghasilan Penanggungnya cukup dan penanggungnya diri sendiri, maka calon penerima beasiswa tersebut layak diberikan beasiswa.
6. Jika calon penerima beasiswa Penghasilan Penanggungnya cukup dan penanggungnya wali, maka calon penerima beasiswa tersebut layak diberikan beasiswa.

7. Jika calon penerima beasiswa Penghasilan Penanggungnya rata-rata dan pekerjaan penanggungnya pekerja tetap dan nilai UN-nya baik dan jumlah tanggungannya banyak, maka calon penerima beasiswa tersebut layak diberikan beasiswa.
8. Jika calon penerima beasiswa Penghasilan Penanggungnya rata-rata dan pekerjaan penanggungnya pekerja tetap dan nilai UN-nya baik dan jumlah tanggungannya sedang, maka calon penerima beasiswa tersebut belum layak diberikan beasiswa.
9. Jika calon penerima beasiswa Penghasilan Penanggungnya rata-rata dan pekerjaan penanggungnya pekerja tetap dan nilai UN-nya baik dan jumlah tanggungannya sedikit, maka calon penerima beasiswa tersebut belum layak diberikan beasiswa.
10. Jika calon penerima beasiswa Penghasilan Penanggungnya rata-rata dan pekerjaan penanggungnya pekerja tetap dan nilai UN-nya cukup maka calon penerima beasiswa tersebut belum layak diberikan beasiswa.
11. Jika calon penerima beasiswa Penghasilan Penanggungnya rata-rata dan pekerjaan penanggungnya pekerja tetap dan nilai UN-nya kurang maka calon penerima beasiswa tersebut belum layak diberikan beasiswa.
12. Jika calon penerima beasiswa Penghasilan Penanggungnya rata-rata dan pekerjaan penanggungnya pekerja tidak tetap, maka calon penerima beasiswa tersebut layak diberikan beasiswa.
13. Jika calon penerima beasiswa Penghasilan Penanggungnya tinggi, maka calon penerima beasiswa tersebut belum layak diberikan beasiswa.
14. Jika calon penerima beasiswa Penghasilan Penanggungnya rendah, maka calon penerima beasiswa tersebut layak diberikan beasiswa.

4. Kesimpulan

Penelitian ini digunakan untuk mendukung keputusan pemberian beasiswa dengan menggunakan algoritma C4.5 yang dapat membantu pemberian beasiswa bagi mahasiswa baru STMIK Bina Bangsa. Dalam penelitian ini didapat kriteria penentu utama beasiswa yaitu penghasilan penanggung, jika penghasilan penanggungnya rendah maka calon penerima beasiswa tersebut layak untuk diberikan beasiswa. Jika penghasilan penanggungnya tinggi, maka calon penerima beasiswa tersebut belum layak diberikan beasiswa. Jika penghasilan penanggungnya cukup atau dalam kategori rata-rata, masih perlu melihat atribut lain untuk menentukan apakah calon penerima beasiswa tersebut layak atau belum layak untuk diberikan beasiswa berdasarkan hasil dan pembahasan yang telah dilakukan pada penelitian ini.

Daftar Pustaka

- [1] Rainer, dedi. 2017. Pengertian Beasiswa, Tujuan, Syarat, Manfaat, Jenis, Contoh Terlengkap. <http://www.studinews.co.id>
- [2] Muh. Safri Juliardi, Ristu saptono, dan Denis Eka Cahyani. 2017 Universitas Sebelas Maret Bidikmisi Applicant's Classification Using C4.5 Algorithm. ITSMART: Jurnal Ilmiah Teknologi dan Informasi Vol. 6, No. 1.
- [3] Jose Augusto Duarte Guterres. 2015. Kelayakan Algoritma C4.5 Sebagai Pendukung Keputusan Dalam Pengajuan Penerima Beasiswa. Seminar Nasional Teknologi Informasi.

- [4] 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.
- [5] Tabrani, Muhamad. 2016. Klasifikasi Penerima Beasiswa Kopertis dengan Menggunakan Algoritma C.45. *Jurnal Pilar Nusa Mandiri* Vol.XII, No. 1, Maret 2016. ISSN 1978-1946.
- [6] Merdekawati, Agustiena. 2018. Sistem Pendukung Keputusan Penerimaan Beasiswa menggunakan Algoritma C4.5 (Studi Kasus : Baitul Maal). *Jurnal Pendidikan Teknologi dan Kejuruan* Vol. 15, No. 1, Januari 2018, Hal: 123. P-ISSN: 0216-3241; E-ISSN: 2541-0652.
- [7] Yutanto, Hariadi dan Nurcholis Setiawan. 2018. Penerapan Data Mining Sebagai Model Seleksi Penerima Beasiswa Penuh (Studi Kasus: STIE Perbanas Surabaya). *Jurnal Link* Vol. 27, No. 1, Februari 2018. ISSN 1858-4667