

ANALISIS FAKTOR -FAKTOR YANG MEMPENGARUHI MAHASISWA DALAM MENGAMBIL BEASISWA DENGAN MENGGUNAKAN ALGORITMA C4.5

Darsono Nababan¹, Alvin Wijaya², William³

¹²³Program Studi Sistem Informasi, Universitas Pelita Harapan Medan

e-mail: *¹darsono.nababan@uph.edu , ² awijaya692@uph.edu, ³ williamphiel@uph.edu,

Abstrak

Pendidikan merupakan salah satu kebutuhan yang sangat penting di dalam masyarakat. Melalui Pendidikan pola pikir dan kualitas hidup seseorang akan meningkat. Untuk mendapatkan Pendidikan yang berkualitas maka dibutuhkan juga biaya yang sangat tinggi saat ini, untuk itu mahasiswa/mahasiswi banyak yang tertarik untuk mengambil beasiswa untuk dapat meringankan beban tersebut. Dengan memanfaatkan pengisian survei untuk mengukur factor-faktor apa saja yang mempengaruhi mahasiswa dalam mengambil beasiswa. Melalui hasil analisis tersebut diharapkan dapat menghasilkan faktor apa yang paling mendekati sehingga mahasiswa mengambil beasiswa. Pengolahan data mining dilakukan dengan menggunakan Algoritma C4.5 untuk membentuk pohon keputusan. Adapun perangkat lunak yang menerapkan Algoritma C4.5 ini adalah Rapid Miner.

Kata kunci—Pendidikan, Beasiswa, Algoritma C4.5, Rapid Miner

Abstract

Education is one of the most important needs in society. Through Education the mindset and quality of one's life will increase. To get quality education, a very high cost is needed at this time, for students, many students are interested in taking scholarships to ease the burden. By utilizing filling surveys to measure what factors influence students in taking scholarships. Through the results of the analysis, it is expected to produce the closest factors so that students take scholarships. Data mining processing is done by using C4.5 Algorithm to form decision trees. The software that applies the C4.5 Algorithm is Rapid Miner.

Keywords— Education, Scholarship, C4.5 Algorithms, Rapid Miner

I. PENDAHULUAN

Sumber daya manusia yang bermutu hanya dapat diwujudkan dengan pendidikan yang bermutu. Pendidikan bermutu adalah pendidikan yang mampu mengembangkan potensi-potensi positif yang terpendam dalam diri siswa didik.

Dengan pendidikan bermutu, pendidikan menghasilkan tenaga-tenaga muda potensial yang tangguh dan siap bersaing dalam masyarakat dunia. Namun pendidikan di Indonesia pada saat ini tergolong rendah dibanding dengan negara-negara lain. Hal tersebut terbukti dengan adanya data yang berasal dari UNESCO dan Balitbang pada tahun 1996-2000. Pada tahun 2008 pendidikan di Indonesia berada diperingkat ke-69 tingkat dunia. Abad 21 merupakan abad pengetahuan oleh karena itu sudah seharusnya pendidikan di Indonesia me-langkah seiring dengan tuntutan jaman agar bangsa Indonesia tidak terlindas oleh jaman akibat ketidak berdayaannya.

Mencerdaskan kehidupan bangsa itu mempunyai 3 komponen yang mem-punyai arti yang sangat penting yaitu (1) cerdas, Cerdas disini bukan berarti hafal seluruh mata pelajaran, yang bingung saat ditanya bagaimana menciptakan solusi bagi kehidupan nyata. Namun yang dimakisud cerdas disini adalah memiliki ilmu yang dapat digunakan untuk menyelesaikan masa-lah yang nyata, kreatif serta inovatif. (2) Hidup, Hidup itu memiliki filosofi untuk menghargai kehidupan dan melakukan hal yang baik untuk kehidupan itu sendiri. Filosofi hidup ini sangat erat akan makna individualisme yang mempunyai arti mengangkat kehidupan seseorang, me-manusiakan seorang manusia, mem-berikannya makna kehidupan berupa semangat, nilai moral dan tujuan hidup. (3) Bangsa, Manusia selain sebagai sosok individu, dia juga sebagai makhluk sosial, dimana antar manusia

saling membutuhkan satu sama lain. Kewajiban sebagai individu yaitu untuk menyebarkan pengetahuannya kepada masyarakat, berusaha meningkatkan derajat kemuliaan masyarakat se-kitarnya dan berperan aktif dalam dinamika masyarakat. Yang dimaksud masyarakat disini adalah identitas bangsa yang menjadi ciri suatu masyarakat.

Namun tujuan pendidikan diatas yang mempunyai arti sangat penting bagi kelangsungan pendidikan di Indonesia belum tercapai secara optimal atau sepenuhnya, sehingga kualitas pendidikan di Indonesia saat ini dalam kategori rendah, hal ini dibuktikan berdasarkan data dari UNESCO (2000) tentang peringkat indeks pengembangan manusia yaitu komposisi dari peringkat pencapaian pendidikan, kesehatan dan penghasilan per kepala yang menunjukkan bahwa indeks pen-gembangan masyarakat Indonesia men-galami penurunan. Dan salah satu faktor rendahnya pendidikan di Indonesia yaitu mahalnya biaya pendidikan, mahalnya biaya pendidikan membuat masyarakat miskin lebih memilih untuk tidak sekolah. Semakin mahalnya pendidikan di Indonesia tidak terlepas dari kebijakan pemerintah yang menerapkan MBS (Manajemen Ber-basis Sekolah). Pada realitanya MBS di Indonesia lebih dimaknai sebagai suatu usaha untuk melakukan mobolisasi dana. Oleh karena itu, komite sekolah/dewan pendidikan sebagai organ MBS memiliki syarat adanya unsur pengusaha, Oleh karena itu untuk membantu masyarakat kurang mampu, beberapa per-kuliahan swasta / negeri memberikan beasiswa, beasiswa yang diberikan juga beragam ada yang tidak perlu membayar sama, sesuai dengan prestasi yang dicapai oleh mahasiswa tersebut. Dalam menganalisa hal tersebut di atas, salah satu solusi yang dapat digunakan adalah data mining mengenai faktor-faktor yang mem-pengaruhi

mahasiswa untuk men-gambil beasiswa dengan menggunakan algoritma C4.5.

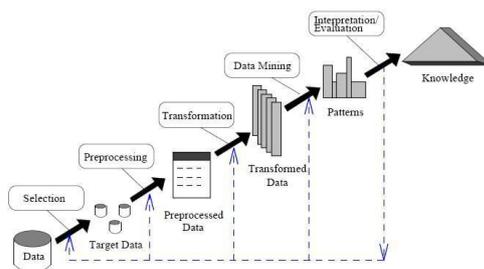
Beberapa penelitian yang sudah dilakukan dengan metode yang sama adalah dalam membantu menentukan persediaan stok yang dapat memenuhi permintaan retailer serta mengatasi penimbunan stok yang berlebih dengan menggunakan teknik data mining klasifikasi dengan metode decision tree. Dengan menggunakan metode ini data set yang didapatkan akan di olah sehingga menjadi informasi yang tersedia terstruktur ke arah pohon keputusan dan mendapatkan pola terbaik untuk menentukan stok yang diperlukan retailer perusahaan dapat menentukan prediksi jumlah persediaan stok barang. [1]

II. TINJAUAN PUSTAKA

2.1 Data Mining

Merupakan proses penggalian dan pertambahan pengetahuan dari sejumlah data yang besar, database atau repository database lainnya. Tujuan utama dari penambangan data ini untuk menemukan pengetahuan baru yang tersembunyi dari database tersebut. Karena Data Mining adalah suatu rangkaian proses yang dapat dibagi menjadi beberapa tahap. Tahap-tahap tersebut bersifat interaktif di mana pemakai terlibat langsung atau dengan perantara knowledge base. [2]

Tahap-tahap ini diilustrasikan pada gambar berikut :



Gambar 1. Tahapan Data Mining

2.2 Pohon Keputusan (Decision Tree)

Pohon keputusan menggunakan representasi struktur pohon (*tree*) di mana setiap node merepresentasikan atribut, cabangnya merepresentasikan nilai dari atribut dan daun merepresentasikan kelas. Node yang paling atas dari pohon keputusan disebut sebagai root. Salah satu keuntungan yang signifikan dari pohon keputusan adalah kenyataan bahwa pengetahuan dapat diekstrak dan direpresentasikan dalam bentuk klasifikasi aturan if-then. Pohon keputusan merupakan metode klasifikasi yang paling populer digunakan. Selain karena pembangunannya relatif cepat, hasil dari model yang dibangun mudah untuk dipahami. Pohon keputusan bekerja dengan membentuk pohon keputusan yang dapat disimpulkan aturan-aturan klasifikasi tertentu, salah satunya adalah C4.5.[3]

2.3 Algoritma C4.5

Algoritma C4.5 adalah algoritma yang sudah banyak dikenal dan digunakan untuk klasifikasi data yang memiliki atribut-atribut numerik dan kategorial. Hasil dari proses klasifikasi yang berupa aturan-aturan dapat digunakan untuk memprediksi nilai atribut bertipe diskret dari record yang baru. Algoritma C4.5 sendiri merupakan pengembangan dari algoritma ID3, dimana pengembangan dilakukan dalam hal, bisa mengatasi missing data, bisa mengatasi data kontinu dan pruning. [4]

Secara umum algoritma C4.5 untuk membangun pohon keputusan adalah sebagai berikut :

1. Pilih atribut sebagai akar.
2. Buat cabang untuk tiap-tiap nilai.
3. Bagi kasus dalam cabang.
4. Ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama.

Untuk memilih atribut akar, didasarkan pada nilai gain tertinggi dari atribut-atribut

yang ada. Untuk menghitung gain digunakan rumus seperti yang tertera dalam persamaan berikut:

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

Dimana :

S : himpunan kasus

A : atribut

N : jumlah partisi atribut A

|Si|: jumlah kasus pada partisi ke-i

|Sr| : jumlah kasus dalam S

III. METODOLOGI PENELITIAN

Metodologi yang digunakan dalam penelitian ini adalah dengan dengan menerapkan langkah -langkah berikut :



Gambar 2. Kerangka Metode Penelitian

1. Pengumpulan Data

Data yang digunakan adalah data mahasiswa baru angkatan 2018 melalui survey dengan kriteria jenis kelamin, lulusan sekolah, tingkat kondisi ekonomi, tempat tinggal, status penerima beasiswa.

2. Analisis Data

Dalam tahap analisis dilakukan proses pengelompokan kedalam bentuk tabel berdasarkan hasil survey.

3. Uji Data

Pengujian data dilakukan dengan menentukan pohon keputusan dengan menggunakan algoritma C.45

IV. HASIL DAN PEMBAHASAN

Analisa data survei faktor yang mempengaruhi mahasiswa mengambil beasiswa dilakukan setelah data terkumpul dan sesuai dengan kebutuhan sistem ini. Oleh sebab itu, untuk menghasilkan kesimpulan berdasarkan pohon keputusan pada analisa data ini, diperlukan data produk yang telah dibeli customer. Analisa data tersebut dilakukan berdasarkan teknik algoritma C4.5 dengan beberapa langkah-langkah yang sudah ditentukan, berikut adalah tabel data hasil survei.

Tabel 1. Kelompok Data Berdasarkan Hasil Survey

Nama	Jenis Kelamin	Lulusan Sekolah	Tingkat Ekonomi	Tinggal	Meriman a Beasiswa
Adeline	Perempuan	SMA	Rendah	Merantau	Ya
Alex	Laki-Laki	SMA	Tinggi	Orang Tua	Ya
Ali	Laki-Laki	SMK	Tinggi	Wali	Ya
Andre	Laki-Laki	SMA	Sedang	Orang Tua	Ya
Andrew	Laki-Laki	SMA	Rendah	Wali	Ya
Andy	Laki-Laki	SMA	Rendah	Orang Tua	Ya
Angel	Perempuan	SMA	Rendah	Merantau	Ya
Angelo	Laki-Laki	SMA	Rendah	Wali	Ya
Angelo	Laki-Laki	SMA	Rendah	Orang Tua	Ya
Anto	Laki-Laki	SMA	Rendah	Merantau	Ya
Asmidar	Perempuan	SMA	Rendah	Orang Tua	Tidak
Aspin	Laki-Laki	SMK	Rendah	Merantau	Ya
Augustambunan	Laki-Laki	SMK	Rendah	Merantau	Tidak
Axel	Laki-Laki	SMA	Sedang	Merantau	Ya
Bahagia Selalu	Laki-Laki	SMA	Sedang	Merantau	Ya
Bakri	Laki-Laki	SMK	Rendah	Orang Tua	Ya

Basuki	Laki-Laki	SMA	Rendah	Merantau	Tidak
Buala	Laki-Laki	SMA	Rendah	Merantau	Tidak
Budi	Laki-Laki	SMK	Tinggi	Merantau	Tidak
Burhanudin	Laki-Laki	SMK	Rendah	Merantau	Tidak
Chris	Laki-Laki	SMK	Rendah	Orang Tua	Ya
Chrislie	Perempuan	SMA	Rendah	Merantau	Ya
Cindy	Perempuan	SMA	Rendah	Orang Tua	Tidak
Citra	Perempuan	SMA	Rendah	Merantau	Tidak
David	Laki-Laki	SMK	Tinggi	Orang Tua	Ya
Dedi	Laki-Laki	SMK	Sedang	Merantau	Tidak
Dedy	Laki-Laki	SMK	Rendah	Wali	Ya
Delvillia	Perempuan	SMA	Sedang	Orang Tua	Tidak
Denise	Laki-Laki	SMA	Rendah	Merantau	Ya
Deny	Laki-Laki	SMA	Rendah	Merantau	Ya
Desy	Perempuan	SMA	Rendah	Wali	Ya
Digo	Laki-Laki	SMA	Rendah	Wali	Ya
Dion	Laki-Laki	SMA	Tinggi	Orang Tua	Ya
Doni	Laki-Laki	SMA	Rendah	Orang Tua	Ya
Duma	Perempuan	SMA	Rendah	Merantau	Ya
Elisabeth	Perempuan	SMK	Sedang	Orang Tua	Ya
Engki	Laki-Laki	SMA	Rendah	Wali	Tidak
Fanny	Perempuan	SMA	Tinggi	Orang Tua	Tidak
Frans	Laki-Laki	SMA	Rendah	Wali	Ya
Harris	Laki-Laki	SMA	Sedang	Orang Tua	Ya
Hendrik	Laki-Laki	SMA	Rendah	Merantau	Ya
Hendro	Laki-Laki	SMK	Sedang	Wali	Ya
Henny	Perempuan	SMA	Rendah	Merantau	Ya
Heri	Laki-Laki	SMA	Sedang	Orang Tua	Tidak

Immanuel	Laki-Laki	SMK	Tinggi	Wali	Tidak
Intan	Perempuan	SMA	Rendah	Merantau	Ya
Ioana	Perempuan	SMA	Rendah	Orang Tua	Ya
Irwanto	Laki-Laki	SMK	Rendah	Orang Tua	Tidak
Jeff	Laki-Laki	SMA	Rendah	Orang Tua	Ya
Johan	Laki-Laki	SMK	Sedang	Merantau	Tidak
Johnny	Laki-Laki	SMA	Tinggi	Orang Tua	Tidak
Joko	Laki-Laki	SMK	Sedang	Merantau	Tidak
Jonathan	Laki-Laki	SMA	Tinggi	Wali	Ya
Joni	Laki-Laki	SMK	Rendah	Wali	Ya
Josceylene	Perempuan	SMA	Rendah	Merantau	Ya
Junaidi	Laki-Laki	SMA	Sedang	Orang Tua	Tidak
Kifli	Laki-Laki	SMA	Rendah	Merantau	Ya
Kiki	Perempuan	SMA	Rendah	Orang Tua	Ya
Leli	Perempuan	SMA	Rendah	Wali	Ya
Leo	Laki-Laki	SMA	Rendah	Orang Tua	Ya
Leo	Laki-Laki	SMK	Sedang	Wali	Ya
Lesley	Perempuan	SMK	Rendah	Merantau	Ya
Lika	Perempuan	SMA	Tinggi	Orang Tua	Ya
Lisa	Perempuan	SMK	Sedang	Merantau	Tidak
Lukman	Laki-Laki	SMK	Sedang	Merantau	Tidak
Maia	Perempuan	SMA	Sedang	Orang Tua	Tidak
Margareth	Perempuan	SMA	Sedang	Orang Tua	Ya
Milea	Perempuan	SMA	Rendah	Merantau	Ya
Miswanto	Laki-Laki	SMA	Rendah	Orang Tua	Tidak
Murni	Perempuan	SMK	Sedang	Merantau	Ya
Nabila	Perempuan	SMA	Tinggi	Wali	Tidak
Natasya	Perempuan	SMK	Sedang	Merantau	Ya

Nicholas	Laki-Laki	SMA	Sedang	Orang Tua	Ya
Nina	Perempuan	SMK	Sedang	Orang Tua	Tidak
Olivia	Perempuan	SMA	Sedang	Orang Tua	Ya
Ordo	Laki-Laki	SMA	Rendah	Orang Tua	Ya
Putri	Perempuan	SMA	Sedang	Merantau	Ya
Radja	Laki-Laki	SMA	Sedang	Orang Tua	Tidak
Rahmad	Laki-Laki	SMK	Rendah	Orang Tua	Tidak
Revindo	Laki-Laki	SMA	Rendah	Merantau	Ya
Ronald	Laki-Laki	SMK	Tinggi	Merantau	Ya
Rosiana	Perempuan	SMA	Tinggi	Merantau	Tidak
Roy	Laki-Laki	SMA	Sedang	Orang Tua	Ya
Shelly	Perempuan	SMK	Sedang	Wali	Ya
Sheren	Perempuan	SMA	Tinggi	Wali	Tidak
Sri	Perempuan	SMA	Sedang	Orang Tua	Ya
Stephanie	Perempuan	SMA	Sedang	Wali	Ya
Suci	Perempuan	SMK	Rendah	Orang Tua	Ya
Tanto	Laki-Laki	SMA	Rendah	Merantau	Ya
Valencia	Perempuan	SMA	Tinggi	Orang Tua	Tidak
Vanessa	Perempuan	SMA	Rendah	Wali	Ya
Vicky	Laki-Laki	SMA	Sedang	Orang Tua	Tidak
Vivi	Perempuan	SMA	Tinggi	Wali	Ya
Wanda	Perempuan	SMK	Sedang	Merantau	Tidak
Wandra	Laki-Laki	SMK	Sedang	Orang Tua	Ya
Wandy	Laki-Laki	SMA	Sedang	Merantau	Tidak
William	Laki-Laki	SMA	Tinggi	Orang Tua	Ya
Windy	Perempuan	SMK	Sedang	Orang Tua	Ya
Yanti	Perempuan	SMA	Sedang	Merantau	Ya
Yoga	Laki-Laki	SMA	Sedang	Orang Tua	Tidak

Yosce	Perempuan	SMA	Rendah	Merantau	Ya
Yudhi	Laki-Laki	SMK	Rendah	Orang Tua	Ya
Yuli	Perempuan	SMA	Rendah	Merantau	Ya
Yusuf	Laki-Laki	SMK	Rendah	Orang Tua	Ya
Zulfan	Laki-Laki	SMA	Rendah	Merantau	Ya

Proses selanjutnya adalah menentukan nilai atribut sebagai akar dan menghitung nilai informasi gain atribut. Untuk memilih atribut sebagai akar, didasarkan pada nilai gain tertinggi dari atribut-atribut yang ada dengan menghitung nilai entropy.

Setelah nilai entropy dari setiap atribut telah didapatkan dan nilai gain dari atribut juga sudah dihitung, selanjutnya hasil dari perhitungan tersebut dimasukkan ke dalam table, berikut adalah tahapan perhitungan atribut yang diambil sebagai sample.

Tahap 1 : Tentukan nilai entropy total menggunakan tabel terakhir yaitu “penerima beasiswa”, dengan menggunakan rumus dibawah :

$$\text{Entropy}(S) \equiv - p_+ \log_2 p_+ - p_- \log_2 p_-$$

Dimana :

S : data sample yang digunakan.

P+ : jumlah yang bersolusi positif (mendukung) pada data sample untuk kriteria tertentu.

P- : jumlah yang bersolusi negatif (tidak mendukung) pada data sample untuk kriteria tertentu.

Nilai Entropy Total Mahasiswa yang menerima beasiswa

Total Data : 105

Total Diterima : 71

Total Ditolak : 34

$$\begin{aligned} \text{Entropy T} &= \left[\begin{aligned} &(-71/105) \\ &\log_2(71/105) \end{aligned} \right] + \left[\begin{aligned} &(- \\ &(34/105) \log_2(34/105) \end{aligned} \right) \\ &= 0,382 + 0,527 \\ &= \mathbf{0,909} \end{aligned}$$

Tahap 2, Tentukan nilai gain pada table “jenis kelamin”, dengan menggunakan rumus dibawah :

$$\text{Entropy(S)} \equiv - p_+ \log_2 p_+ - p_- \log_2 p_-$$

Dimana:

- S : data sample yang digunakan.
P+ : jumlah yang bersolusi positif (men-dukung) pada data sample untuk kriteria tertentu.
P- : jumlah yang bersolusi negatif(tidak mendukung) pada data sample untuk kriteria tertentu.

Nilai gain berdasarkan “Faktor Berdasarkan Jenis Kelamin”

Total Data : 105
Laki-Laki : 62
Perempuan : 43

Entropy (Laki-Laki):
Total: 62
Diterima: 41
Ditolak : 21

$$\begin{aligned} \text{Entropy (Jenis kelamin(Laki-Laki))} \\ &= \left(\begin{aligned} &-(41/62) \log_2(41/62) \end{aligned} \right) + \left(\begin{aligned} &-(21/62) \\ &\log_2(21/62) \end{aligned} \right) \\ &= 0,395 + 0,529 \\ &= 0,924 \end{aligned}$$

Entropy (Perempuan):
Total: 43
Diterima: 30
Ditolak : 13

$$\begin{aligned} \text{Entropy (Jenis kelamin(Perempuan))} \\ &= \left(\begin{aligned} &-(30/43) \log_2(30/43) \end{aligned} \right) + \left(\begin{aligned} &-(13/43) \\ &\log_2(13/43) \end{aligned} \right) \\ &= 0,362 + 0,522 \\ &= 0,884 \end{aligned}$$

Setelah mendapat Entropy dari data jenis kelamin maka kita akan mencari gain dari jenis kelamin

$$\begin{aligned} \text{Gain (Jenis Kelamin)} \\ &= 0,909 - \left(\begin{aligned} &((62/105) \times 0,924) \end{aligned} \right) + \left(\begin{aligned} &((43/105) \times \\ &0,884) \end{aligned} \right) \\ &= 0,909 - 0,908 \\ &= \mathbf{0,001} \end{aligned}$$

Maka hasil dari gain jenis kelamin yaitu: **0,001**

Nilai Gain berdasarkan “Faktor Lulusan sekolah”

Total Data: 105
SMA: 72
SMK: 33

$$\begin{aligned} \text{Entropy (Lulusan sekolah(SMA))} \\ &= \left(\begin{aligned} &-(51/72) \log_2(51/72) \end{aligned} \right) + \left(\begin{aligned} &-(21/72) \\ &\log_2(21/72) \end{aligned} \right) \\ &= 0,353 + 0,518 \\ &= 0,871 \end{aligned}$$

$$\begin{aligned} \text{Entropy(Lulusan sekolah(SMK))} \\ &= \left(\begin{aligned} &-(20/33) \log_2(20/33) \end{aligned} \right) + \left(\begin{aligned} &-(13/33) \\ &\log_2(13/33) \end{aligned} \right) \\ &= 0,438 + 0,529 \\ &= 0,967 \end{aligned}$$

$$\begin{aligned} \text{Gain (Lulusan sekolah)} \\ &= 0,909 - \left(\begin{aligned} &((72/105) \times 0,872) \end{aligned} \right) + \left(\begin{aligned} &((33/105) \times \\ &0,967) \end{aligned} \right) \\ &= 0,909 - 0,902 \\ &= \mathbf{0,007} \end{aligned}$$

Maka hasil dari gain lulusan sekolah yaitu: **0,007**

Nilai Gain berdasarkan “Faktor Tingkat Ekonomi”

Total Data: 105
Rendah : 53
Sedang : 35
Tinggi : 17

$$\begin{aligned} \text{Entropy (Tingkat Ekonomi(Rendah))} \\ &= \left(\begin{aligned} &-(42/53) \log_2(42/53) \end{aligned} \right) + \left(\begin{aligned} &-(11/53) \\ &\log_2(11/53) \end{aligned} \right) \\ &= 0,266 + 0,471 \end{aligned}$$

$$= 0,737$$

$$\begin{aligned} &\text{Entropy (Tingkat Ekonomi(Sedang))} \\ &= (-\frac{20}{35} \log_2(\frac{20}{35})) + (-\frac{15}{35} \log_2(\frac{15}{35})) \\ &= 0,461 + 0,524 \\ &= 0,985 \end{aligned}$$

$$\begin{aligned} &\text{Entropy (Tingkat Ekonomi(Tinggi))} \\ &= (-\frac{9}{17} \log_2(\frac{9}{17})) + (-\frac{8}{17} \log_2(\frac{8}{17})) \\ &= 0,486 + 0,512 \\ &= 0,998 \end{aligned}$$

$$\begin{aligned} &\text{Gain (Tingkat Ekonomi)} \\ &= 0,909 - ((\frac{53}{105}) \times 0,737 + (\frac{35}{105}) \times 0,985 + (\frac{17}{105}) \times 0,998) \\ &= 0,909 - 0,862 \\ &= \mathbf{0,047} \end{aligned}$$

Maka hasil dari gain tingkat ekonomi yaitu: **0,047**

Nilai Gain berdasarkan “Faktor Tempat Tinggal”

Total Data: 105

Merantau : 41

Orangtua : 44

Wali : 20

$$\begin{aligned} &\text{Entropy (Tempat Tinggal(Merantau))} \\ &= (-\frac{27}{41} \log_2(\frac{27}{41})) + (-\frac{14}{41} \log_2(\frac{14}{41})) \\ &= 0,397 + 0,529 \\ &= 0,926 \end{aligned}$$

$$\begin{aligned} &\text{Entropy (Tempat Tinggal(Orangtua))} \\ &= (-\frac{28}{44} \log_2(\frac{28}{44})) + (-\frac{16}{44} \log_2(\frac{16}{44})) \\ &= 0,415 + 0,531 \\ &= 0,946 \end{aligned}$$

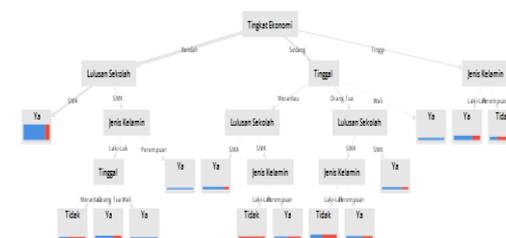
$$\begin{aligned} &\text{Entropy (Tempat Tinggal(Wali))} \\ &= (-\frac{16}{20} \log_2(\frac{16}{20})) + (-\frac{4}{20} \log_2(\frac{4}{20})) \\ &= 0,258 + 0,464 \\ &= 0,722 \end{aligned}$$

$$\begin{aligned} &\text{Gain (Tempat Tinggal)} \\ &= 0,909 - ((\frac{41}{105}) \times 0,926 + (\frac{44}{105}) \times 0,946 + (\frac{20}{105}) \times 0,722) \\ &= 0,909 - 0,896 \\ &= \mathbf{0,013} \end{aligned}$$

Maka hasil dari gain tempat tinggal yaitu: **0,013**

Dari hasil gain yang di dapat dari masing-masing data dapat disimpulkan bahwa faktor yang paling mempengaruhi mahasiswa untuk mengambil beasiswa yaitu faktor tingkat ekonomi dengan gain tertinggi “0,047”, dan setelah mendapatkan hasil dari gain, akan kita akan coba menggunakan *tools* rapid miner untuk memproses hasil dari perhitungan manual.

Berdasarkan hasil proses pada *tools* rapid miner dapat ditunjukkan hasilnya pada gambar 2 berikut ini :



Gambar 3. Rapid Miner Proses ke-1

Dari hasil pohon keputusan pada gambar 2 dapat dilihat bahwa faktor yang paling mempengaruhi mahasiswa untuk men-gambil beasiswa yaitu faktor tingkat ekonomi.

Maka dari hasil yang didapatkan dari rapid miner dimana Tingkat Ekonomi menjadi faktor utama mahasiswa untuk mengambil beasiswa sama dengan perhitungan manual.

V. KESIMPULAN

Berdasarkan tujuan pembahasan dan implementasi yang telah dilakukan maka dapat diperoleh beberapa kesimpulan yaitu :

1. Algoritma C4.5 dapat digunakan untuk mengenali faktor yang mempengaruhi mahasiswa untuk mengambil beasiswa.
2. Setelah melakukan implementasi data mining dengan menggunakan metode algoritma C4.5, maka karakteristik faktor mahasiswa mengambil beasiswa

telah dapat diketahui mana faktor yang paling mempengaruhi dari hasil survei tersebut, dan diharapkan dari penelitian ini dapat memberikan sasaran yang lebih tepat dari pemberi beasiswa agar orang-orang yang berprestasi dapat diberi kesempatan untuk berkuliah, dan memajukan bangsa kita yaitu Indonesia.

VI DAFTAR PUSTAKA

- [1] D. Nababan and A. V. Tanlim, "Analisis Persediaan Stok Barang Menggunakan Algoritma C 4 . 5 (Studi Kasus CV Harapan Raya)," *J. Ilm. Inform.*, vol. 7, no. 1, pp. 1–10, 2019.
- [2] D. Nababan, "Analisa Algoritma C4 . 5 Dalam Menentukan Faktor Yang Mempengaruhi Munculnya Professional Blogger," *J. Ilm. Inform.*, vol. 7, no. 1, pp. 11–17, 2019.
- [3] Darsono Nababan and Robbi Rahim, "Sistem Pendukung Keputusan Reward Bonus Karyawan Dengan Metode Topsis," *Simetris*, vol. 3, no. 6, pp. 2–6, 2018.
- [4] A. Andriani, "Sistem Pendukung Keputusan Berbasis Decision Tree Dalam Pemberian Beasiswa Studi Kasus : Amik ' Bsi Yogyakarta ,'" *Semin. Nas. Teknol. Inf. dan Komun. 2013 (SENTIKA 2013)*, vol. 2013, no. Sentika, pp. 163–168, 2013.