

PENERAPAN DATA MINING UNTUK KLASIFIKASI SELEKSI  
CALON ANGGOTA HMIF AMIKOM YOGYAKARTA  
MENGUNAKAN  
METODE ALGORITMA DECISION TREE C4.5

Windha Mega Pradnya D<sup>1)</sup>, Mita Pertiwi <sup>2)</sup>, Ervan Febriyanto <sup>3)</sup>, Zian Fahrudy<sup>4)</sup>  
<sup>1,2,3,4</sup>Informatika, Universitas Amikom Yogyakarta

e-mail: <sup>1</sup> [windha@amikom.ac.id](mailto:windha@amikom.ac.id), <sup>2</sup> [mitapertiwi267@gmail.com](mailto:mitapertiwi267@gmail.com), <sup>3</sup>  
[evanfebriyanto054@gmail.com](mailto:evanfebriyanto054@gmail.com), <sup>4</sup> [zianfahrudy@gmail.com](mailto:zianfahrudy@gmail.com)

**Abstract**

*Members of the Amikom HMIF Organizational Committee then have the authority to carry out their vision and mission or main objectives at Amikom Yogyakarta University, one of which is the implementation of recruitment and placement according to requirements and procedures that have been determined by the Chairperson and other authorized members according to the needs of the formation in each organizational unit or body. Many cases have been found that the selection of members from the same year and during the interview.*

*In the case of the initial C4.5 algorithm, in determining whether or not candidates are eligible to use graduation parameters, namely oral and written interviews, enthusiasm and achievement in organization in the history of education that has been followed. Data mining benefits can be implemented in data parts that contain data in large quantities. The technique used in data mining is a decision tree with a C4.5 decision tree. The results of the study used 80 data from prospective members who had registered. Producing experimental and evaluation results that C4.5 decision tree algorithm is accurate in determining whether or not prospective members pass.*

**Keywords**— *Selection of prospective members, classifications, Decision Tree C4.5 algorithm*

PENDAHULUAN

HMIF merupakan himpunan jurusan sarjana informatika yang berperan sebagai penyelenggara seleksi calon anggota baru serta penempatan struktur keanggotaan di HMIF. Salah satu tugas dan wewenang HMIF atas pelaksanaan kegiatan dalam organisasi untuk mendukung perkembangan serta kemajuan universitas amikom Yogyakarta, menampung semua aspirasi dari mahasiswa sarjana informatika, membantu organisasi lain untuk mempererat silaturahmi antar organisasi serta membantu mahasiswa dalam belajar sebelum uts maupun uas dengan tema belajar bareng HMIF.

Melihat banyaknya mahasiswa yang mendaftar di HMIF dari pihak organisasi telah menentukan batas penerimaan anggota baru yaitu 40 mahasiswa dalam masa 1 periode itu 1 tahun atau 2 semester. Oleh sebab itu panitia mengadakan tes penerimaan calon anggota baru untuk mengetahui tingkat minat calon anggota, prestasi dan keuletan dalam berorganisasi untuk menghindari keluarnya salah satu anggota sebelum masa periodenya selesai, partisipasi dalam kegiatan HMIF kurang serta masalah individu antar teman lain dalam organisasi. Sebagai langkah dan solusi yang di ajukan sebagai penyelesaian terhadap uraian diatas maka penulis tertarik untuk mengklasifikasi suatu permasalahan sistem data mining yang mudah digunakan dan akurat yaitu mengklasifikasi data mining untuk menampilkan informasi penerimaan seleksi calon anggota baru pada himpunan informatika amikom Yogyakarta menggunakan algoritma decision tree C4.5

Sehingga pelaksana pengadaan calon anggota HMIF dapat berjalan sesuai prosedur yang ditetapkan oleh semua anggota HMIF. Serta dapat diketahui kriteria kriteria apa saja yang dibutuhkan dalam pelaksanaan kegiatan dan seleksi calon anggota baru di HMIF amikom Yogyakarta sebagai bahan klasifikasi

Data mining sendiri memiliki beberapa algoritma salah satunya yaitu decision tree yang merupakan metode klasifikasi yang biasa digunakan karena mudah dan banyak implementasi di berbagai bidang karena sederhana, serta memiliki kemampuan akurasi yang baik untuk mengklasifikasi banyaknya data. Klasifikasi merupakan proses dari mencari suatu himpunan model (fungsi) yang dapat mendeskripsikan dan membedakan kelas-kelas data atau konsep-konsep, dengan tujuan dapat menggunakan model tersebut untuk memprediksi kelas dari suatu objek yang mana kelasnya belum diketahui. [1]

Algoritma C4.5 merupakan algoritma yang banyak di kenal dan sering di gunakan dalam klasifikasi data yang mempunyai atribut numerik dan kategorial. [2] Hasil dari proses klasifikasi yang berupa aturan aturan untuk memprediksi nilai atribut bertipe diskret dari record yang baru. Pengimplementasian metode decision tree C4.5 memiliki tingkat akurasi yang baik untuk solusi pengklasifikasian data dalam jumlah besar dengan teknik data mining menggunakan algoritma C4.5 [3]. Decision tree juga pengembangan dari algoritma ID3, yaitu dapat mengatasi missing data, mengatasi data kontinu dan pruning. Algoritma ini memiliki kelebihan yaitu mudah di mengerti, fleksibel dan dapat divisualisasikan dalam bentuk gambar (pohon keputusan) membuat pohon keputusan dari node atas hingga cabang terakhirnya, dimana atribut paling atas adalah akar dan yang paling bawah adalah daun. Karena metode ini digunakan dalam klasifikasi maka hasil dari olah dataset berupa pengelompokan data ke dalam kelas masing-masing.

## METODE

Pada Penelitian ini data yang digunakan adalah data calon anggota HMIF Universitas Amikom Yogyakarta Dalam penelitian ini akan lakukan dengan langkah-langkah sebagai berikut :

### A. Pengumpulan Data

Dalam penelitian ini, data set yang digunakan adalah tes tertulis, tes wawancara serta prestasi. Penulis mencari informasi dari hasil wawancara dengan ketua himpunan serta wawancara dengan anggota panitia himpunan di dukung dengan beberapa sampel tes penerimaan anggota baru tahun 2017. Data yang kita dapat akan di kumpulkan kemudian disimpan di dalam database dan excel.

### B. Perancangan Metode yang diusulkan

Metode yang digunakan yaitu Algoritma Decision Tree C4.5 untuk mengklasifikasikan calon anggota baru organisasi HMIF.

#### **Tahapan Algoritma Decision Tree C4.5 [3]**

- 1) Siapkan data traning
- 2) Pilih atribut sebagai layar

Rumus menghitung entropy pada algoritma C4.5 [4] :

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

.....(1)

Keterangan :

**S** : Himpunan Kasus

**n** : Jumlah partisi S

**pi** : Proporsi Si terhadap S

Setelah mengetahui nilai entropy dari keseluruhan kasus, maka dilakukan analisis pada setiap atribut dan nilai nilainya kemudian hitung Gain dengan rumus berikut [5] :

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

.....(2)

Keterangan :

**S** : Himpunan Kasus

**A** : Atribut

**n** : Jumlah partisi atribut A

**|Si|** : Proporsi Si terhadap S

**|S|** : Jumlah kasus dalam S

- 3) Buat cabang untuk tiap-tiap nilai
- 4) Ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama
- 5) Pohon Keputusan

### C. Implementasi

Implementasi penelitian dilakukan dengan langkah kerja : mempelajari topik yang di bahas > pengumpulan data dari wawancara maupun hasil tes tertulis dari calon anggota > processing perhitungan presentase syarat kelulusan > ekstrasi fitur > training > testing > evaluasi. Dengan cara ini maka akan bermanfaat untuk membantu panitia dalam menilai serta memilih anggota yang akan diterima tanpa ada gangguan dari luar seperti panitia meloloskan calon anggota yang disukainya' jadi untuk mengurangi dampak negatif kecurangan itu terjadi. Selanjutnya dataset akan dilakukan pemilahan berdasarkan kelas masing masing sesuai kemampuan dari hasil testing calon anggota. Untuk menentukan klasifikasi dengan menggunakan Decision Tree C4.5 dan pohon keputusan.

### HASIL DAN PEMBAHASAN

Analisis data sampel yang kita gunakan yaitu 80 data sampel calon anggota organisasi HMIF yang akan diklasifikasikan menggunakan Decision Tree C4.5. Tools yang digunakan adalah Rapidminer v.6.5.2 untuk mengetahui pengklasifikasian dengan hasil pohon keputusan yang akurat. Berikut data yang kami gunakan dengan parameter sebagai berikut .:

TABLE I. TABEL DATA SAMPLE

NOMOR PENDAFTARAN	ANGKATAN	JENIS KELAMIN	NILAI	HASIL
001	2017	LAKI-LAKI	RENDAH	TIDAK DITERIMA
002	2017	LAKI-LAKI	RENDAH	TIDAK DITERIMA
003	2016	PEREMPUAN	RENDAH	TIDAK DITERIMA
004	2017	PEREMPUAN	RENDAH	TIDAK DITERIMA
005	2017	LAKI-LAKI	RENDAH	TIDAK DITERIMA
006	2017	PEREMPUAN	TINGGI	DITERIMA
007	2017	LAKI-LAKI	RENDAH	TIDAK DITERIMA
008	2017	LAKI-LAKI	RENDAH	TIDAK DITERIMA
009	2017	LAKI-LAKI	RENDAH	TIDAK DITERIMA
010	2017	LAKI-LAKI	RENDAH	TIDAK DITERIMA
011	2017	PEREMPUAN	RENDAH	TIDAK DITERIMA
012	2017	LAKI-LAKI	SEDANG	TIDAK DITERIMA
013	2017	PEREMPUAN	SEDANG	TIDAK DITERIMA
014	2016	PEREMPUAN	SEDANG	TIDAK DITERIMA
015	2017	LAKI-LAKI	SEDANG	TIDAK DITERIMA
016	2016	PEREMPUAN	SEDANG	TIDAK DITERIMA
017	2016	LAKI-LAKI	SEDANG	TIDAK DITERIMA
018	2016	LAKI-LAKI	SEDANG	TIDAK DITERIMA
019	2016	LAKI-LAKI	SEDANG	TIDAK DITERIMA
020	2016	LAKI-LAKI	SEDANG	TIDAK DITERIMA
021	2016	LAKI-LAKI	SEDANG	TIDAK DITERIMA
022	2017	LAKI-LAKI	SEDANG	TIDAK DITERIMA
023	2017	LAKI-LAKI	SEDANG	TIDAK DITERIMA
024	2016	LAKI-LAKI	SEDANG	TIDAK DITERIMA
025	2016	PEREMPUAN	TINGGI	DITERIMA
026	2016	PEREMPUAN	TINGGI	DITERIMA
027	2017	PEREMPUAN	TINGGI	DITERIMA
028	2017	PEREMPUAN	TINGGI	DITERIMA
029	2016	PEREMPUAN	TINGGI	DITERIMA
030	2017	LAKI-LAKI	TINGGI	DITERIMA
031	2017	LAKI-LAKI	TINGGI	DITERIMA
032	2017	LAKI-LAKI	TINGGI	DITERIMA
033	2016	PEREMPUAN	TINGGI	DITERIMA
034	2017	PEREMPUAN	TINGGI	DITERIMA
035	2016	PEREMPUAN	TINGGI	DITERIMA
036	2016	LAKI-LAKI	TINGGI	DITERIMA
037	2016	LAKI-LAKI	TINGGI	DITERIMA
038	2017	LAKI-LAKI	TINGGI	DITERIMA
039	2017	PEREMPUAN	TINGGI	DITERIMA
040	2017	LAKI-LAKI	SEDANG	TIDAK DITERIMA
041	2017	LAKI-LAKI	SEDANG	TIDAK DITERIMA
042	2016	PEREMPUAN	SEDANG	DITERIMA
043	2017	LAKI-LAKI	SEDANG	DITERIMA
044	2016	PEREMPUAN	TINGGI	DITERIMA

045	2016	LAKI-LAKI	TINGGI	DITERIMA
046	2016	PEREMPUAN	TINGGI	DITERIMA
047	2016	PEREMPUAN	RENDAH	TIDAK DITERIMA
048	2017	LAKI-LAKI	RENDAH	TIDAK DITERIMA
049	2017	LAKI-LAKI	RENDAH	TIDAK DITERIMA
050	2016	LAKI-LAKI	RENDAH	TIDAK DITERIMA
051	2017	LAKI-LAKI	RENDAH	TIDAK DITERIMA
052	2017	LAKI-LAKI	TINGGI	DITERIMA
053	2017	PEREMPUAN	TINGGI	DITERIMA
054	2017	PEREMPUAN	SEDANG	TIDAK DITERIMA
055	2017	LAKI-LAKI	SEDANG	DITERIMA
056	2017	PEREMPUAN	SEDANG	TIDAK DITERIMA
057	2017	LAKI-LAKI	TINGGI	DITERIMA
058	2017	LAKI-LAKI	TINGGI	DITERIMA
059	2017	LAKI-LAKI	TINGGI	DITERIMA
060	2016	PEREMPUAN	TINGGI	DITERIMA
061	2016	PEREMPUAN	SEDANG	TIDAK DITERIMA
062	2016	PEREMPUAN	SEDANG	TIDAK DITERIMA
063	2017	LAKI-LAKI	SEDANG	DITERIMA
064	2017	LAKI-LAKI	TINGGI	DITERIMA
065	2017	LAKI-LAKI	TINGGI	DITERIMA
066	2017	PEREMPUAN	TINGGI	DITERIMA
067	2017	PEREMPUAN	TINGGI	DITERIMA
068	2017	LAKI-LAKI	TINGGI	DITERIMA
069	2017	LAKI-LAKI	RENDAH	TIDAK DITERIMA
070	2017	LAKI-LAKI	RENDAH	TIDAK DITERIMA
071	2017	LAKI-LAKI	RENDAH	TIDAK DITERIMA
072	2017	LAKI-LAKI	RENDAH	TIDAK DITERIMA
073	2017	PEREMPUAN	TINGGI	DITERIMA
074	2016	PEREMPUAN	TINGGI	DITERIMA
075	2016	PEREMPUAN	TINGGI	DITERIMA
076	2016	LAKI-LAKI	TINGGI	DITERIMA
077	2017	LAKI-LAKI	SEDANG	TIDAK DITERIMA
078	2016	PEREMPUAN	SEDANG	DITERIMA
079	2016	LAKI-LAKI	SEDANG	TIDAK DITERIMA
080	2016	PEREMPUAN	SEDANG	DITERIMA

Pengumpulan data calon anggota baru

Syarat kelulusan

Tes tertulis : 10%

Tes wawancara : 30%

LDKMO : 60%

Berdasarkan tabel diatas merupakan data sampel calon anggota yang akan diklasifikasikan menurut karakteristik yang telah ditentukan yaitu tes tertulis, tes wawancara, dan LDKMO. Penilaian tes tertulis dilakukan untuk mengetahui tingkat

pengetahuan calon anggota mengenai Universitas Amikom Yogyakarta dan Himpunan Mahasiswa Informatika (HMIF) Amikom. Sedangkan tes wawancara dilakukan untuk mengetahui kesiapan calon anggota menjadi pengurus HMIF dengan memberikan pertanyaan mengenai loyalitas calon anggota dan kesanggupan calon anggota untuk mengikuti tes terakhir yakni LDKMO. Dan tahap tes seleksi terakhir yakni LDKMO (Latihan Dasar Kepemimpinan dan Manajemen Organisasi) yang merupakan kegiatan wajib yang harus diikuti calon anggota untuk dapat dipertimbangkan menjadi pengurus HMIF. Seperti persyaratan kelulusan diatas yaitu diambil dari persentase tes tulis 10%, tes wawancara 30% dan LDKMO 60%. Calon anggota harus mengikuti ketiga tahap seleksi ini agar klasifikasi nilai bias menjadi maksimal. Kemudian jika terdapat calon anggota memiliki sertifikat berupa piagam prestasi sebelumnya di Sekolah Menengah Pertama maupun Sekolah Menengah Atas maka akan menjadi nilai lebih untuk penilaian calon anggota. Kemudian dari tahap seleksi terakhir yaitu LDKMO dilakukan penilaian sifat individual calon anggota yang meliputi rasa kekeluargaan, sifat profesionalitas, berpikir kritis dan kepedulian antar sesama pengurus.

TABLE II. TABEL CONVERSI DATA

VARIABEL	KETERANGAN	KELOMPOK
PENILAIAN TEORI		
TERTULIS	<80	1
	>80	2
WAWANCARA	<80	1
	>80	2
LDKMO	<80	1
	>80	2
PENILAIAN SIFAT		
KEKELUARGAAN	A	1
	B	2
	C	3
PROFESIONALITAS	A	1
	B	2

	C	3
KRITIS	A	1
	B	2
	C	3
KEPEDULIAN	A	1
	B	2
	C	3
LOYALITAS	A	1
	B	2
	C	3

Tabel II di atas merupakan konversi data untuk penilaian calon anggota dari segi penilaian teori dan penilaian sifat. terdapat tes, yaitu tes tertulis

Dari Tabel 1. Diatas kemudian dilakukan proses menghitung Entropy dan Gain untuk mendapatkan pohon keputusan agar pengklasifikasian calon anggota baru HMIF lebih akurat.

Perhitungan Entropy Akar

### Entropy Total

$$\text{Entropy (total)} = \left(-\frac{40}{80} * \log_2 \left(\frac{40}{80}\right)\right) + \left(-\frac{40}{80} * \log_2 \left(\frac{40}{80}\right)\right) = 0.5 + 0.5 = 1$$

### Entropy Angkatan

$$\begin{aligned} \text{Entropy (2016)} &= \left(-\frac{14}{31} * \log_2 \left(\frac{14}{31}\right)\right) + \left(-\frac{17}{31} * \log_2 \left(\frac{17}{31}\right)\right) = \\ &0.51792836890967741935483870967742 + 0.47530545081870967741935483870968 = \\ &0.9932338197283870967741935483871 \end{aligned}$$

$$\begin{aligned} \text{Entropy (2017)} &= \left(-\frac{26}{49} * \log_2 \left(\frac{26}{49}\right)\right) + \left(-\frac{23}{49} * \log_2 \left(\frac{23}{49}\right)\right) = \\ &0.4851229239840816326530612244898 + 0.51217145767959183673469387755102 = \\ &0.99729438166367346938775510204082 \end{aligned}$$

### Entropy Jenis Kelamin

$$\begin{aligned} \text{Entropy (Laki-Laki)} &= \left(-\frac{29}{47} * \log_2 \left(\frac{29}{47}\right)\right) + \left(-\frac{14}{47} * \log_2 \left(\frac{14}{47}\right)\right) = \\ 0.42982186893510638297872340425532 &+ 0.52045265988085106382978723404255 = \\ 0.95027452881595744680851063829787 & \end{aligned}$$

$$\begin{aligned} \text{Entropy (Perempuan)} &= \left(-\frac{11}{33} * \log_2 \left(\frac{11}{33}\right)\right) + \left(-\frac{22}{33} * \log_2 \left(\frac{22}{33}\right)\right) = \\ 0.52832083356666666666666666666667 &+ 0.38997500048 = \\ 0.91829583404666666666666666666667 & \end{aligned}$$

**Entropy Nilai**

$$\text{Entropy (Tinggi)} = \left(-\frac{0}{34} * \log_2 \left(\frac{0}{34}\right)\right) + \left(-\frac{34}{34} * \log_2 \left(\frac{34}{34}\right)\right) = 0 + 0 = 0$$

$$\begin{aligned} \text{Entropy (Sedang)} &= \left(-\frac{21}{27} * \log_2 \left(\frac{21}{27}\right)\right) + \left(-\frac{6}{27} * \log_2 \left(\frac{6}{27}\right)\right) = \\ 0.28199895062888888888888888888889 &+ 0.48220555586666666666666666666667 = \\ 0.76420450649555555555555555555556 & \end{aligned}$$

$$\text{Entropy (Rendah)} = \left(-\frac{19}{19} * \log_2 \left(\frac{19}{19}\right)\right) + \left(-\frac{0}{19} * \log_2 \left(\frac{0}{19}\right)\right) = 0$$

**Perhitungan Gain Akar**

$$\begin{aligned} \text{Gain (Total, Angkatan)} &= 1 - \\ \left(\left(\frac{31}{80}\right) * 0.9932338197283870967741935483871\right) &+ \left(\frac{49}{80} * \right. \\ \left.0.99729438166367346938775510204082\right) &= \end{aligned}$$

$$1 - (0.38487810514475 + 0.610842808769) = 0.00427908608625$$

$$\begin{aligned} \text{Gain (Total, Jenis Kelamin)} &= 1 - \\ \left(\left(\frac{47}{80}\right) * 0.95027452881595744680851063829787\right) &+ \left(\frac{33}{80} * \right. \\ \left.0.91829583404666666666666666666667\right) &= \end{aligned}$$

$$1 - (0.558286285679375 + 0.37879703154425) = 0.062916682776375$$

$$\begin{aligned} \text{Gain (Total, Nilai)} &= 1 - \\ \left(\left(\frac{34}{80}\right) * 0\right) &+ \left(\frac{27}{80} * 0.76420450649555555555555555555556\right) + \left(\frac{19}{80} * 0\right) = \\ 1 - (0 + 0.25791902094225 + 0) &= 0.74208097905775 \end{aligned}$$

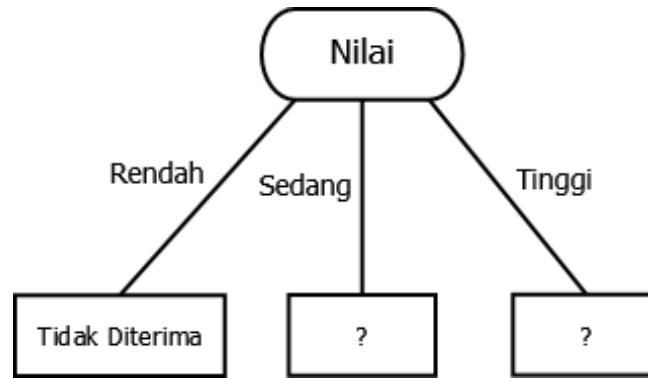
TABLE III. TABEL HASIL PERHITUNGAN ENTROPHY DAN GAIN NODE 1

NODE	KRITERIA	ATRIBUT	JUMLAH KASUS	DITERIMA	TIDAK DITERIMA	ENTROPHY	GAIN
					A		



1	TOTAL		80	40	40	1	
	ANGKATAN						0.004279 08608625
		2016	31	17	14	0.99323381972 8387096774193 5483871	
		2017	49	23	26	0.99729438166 3673469387755 10204082	
	JENIS KELAMIN						0.062916 68277637 5
		LAKI- LAKI	47	18	29	0.95027452881 5957446808510 63829787	
		PEREMPU AN	33	22	11	0.91829583404 6666666666666 66666667	
	NILAI						0.742080 97905775
		TINGGI	34	34	0	0	
		SEDANG	27	6	21	0.76420450649 5555555555555 55555556	
		RENDAH	19	0	19	0	

Tabel III merupakan hasil perhitungan entropy dan gain pada node 1. Dapat diketahui bahwa atribut dengan Gain tertinggi adalah Nilai yaitu sebesar 0.74208097905775. Dengan demikian atribut Nilai dapat menjadi node akar. Dari hasil perhitungan Entrophy dan Gain pada node 1 dapat disimpulkan pohon akar untuk sementara sebagai berikut



Gambar 1. Hasil pohon akar node 1

Gambar 1 adalah Hasil pohon akar pada node 1 yaitu Atribut Nilai dan memiliki 3 cabang, Tinggi, Sedang dan Rendah. Dari ketiga cabang tersebut, nilai cabang atribut Rendah sudah mengklasifikasikan kasus menjadi 1 yaitu keputusan-nya Tidak Diterima, dan cabang atribut Tinggi juga sudah mengklasifikasikan kasus menjadi 1 yaitu keputusannya Diterima, sehingga tidak perlu dilakukan perhitungan lebih lanjut, namun untuk cabang Sedang masih perlu dilakukan perhitungan lagi.

TABLE IV. TABEL ATRIBUT CABANG

NOMOR PENDAFTARAN	ANGKATAN	JENIS KELAMIN	NILAI	HASIL
012	2017	LAKI-LAKI	SEDANG	TIDAK DITERIMA
013	2017	PEREMPUAN	SEDANG	TIDAK DITERIMA
014	2016	PEREMPUAN	SEDANG	TIDAK DITERIMA
015	2017	LAKI-LAKI	SEDANG	TIDAK DITERIMA
016	2016	PEREMPUAN	SEDANG	TIDAK DITERIMA
017	2016	LAKI-LAKI	SEDANG	TIDAK DITERIMA
018	2016	LAKI-LAKI	SEDANG	TIDAK DITERIMA
019	2016	LAKI-LAKI	SEDANG	TIDAK DITERIMA
020	2016	LAKI-LAKI	SEDANG	TIDAK DITERIMA
021	2016	LAKI-LAKI	SEDANG	TIDAK DITERIMA
022	2017	LAKI-LAKI	SEDANG	TIDAK DITERIMA
023	2017	LAKI-LAKI	SEDANG	TIDAK DITERIMA
024	2016	LAKI-LAKI	SEDANG	TIDAK DITERIMA
040	2017	LAKI-LAKI	SEDANG	TIDAK DITERIMA
041	2017	LAKI-LAKI	SEDANG	TIDAK DITERIMA
042	2016	PEREMPUAN	SEDANG	DITERIMA
043	2017	LAKI-LAKI	SEDANG	DITERIMA
054	2017	PEREMPUAN	SEDANG	TIDAK DITERIMA
055	2017	LAKI-LAKI	SEDANG	DITERIMA
056	2017	PEREMPUAN	SEDANG	TIDAK DITERIMA
061	2016	PEREMPUAN	SEDANG	TIDAK DITERIMA
062	2016	PEREMPUAN	SEDANG	TIDAK DITERIMA
063	2017	LAKI-LAKI	SEDANG	DITERIMA
077	2017	LAKI-LAKI	SEDANG	TIDAK DITERIMA
078	2016	PEREMPUAN	SEDANG	DITERIMA
079	2016	LAKI-LAKI	SEDANG	TIDAK DITERIMA
080	2016	PEREMPUAN	SEDANG	DITERIMA

Tabel IV merupakan table atribut cabang yang telah diklasifikasikan dari perhitungan pada node 1 dan akan dilanjutkan untuk perhitungan pada node 1.1

Perhitungan atribut cabang Entropy dan Gain Akar

### Entropy Angkatan

$$\text{Entropy (2016)} = \left(-\frac{11}{14} * \log_2 \left(\frac{11}{14}\right)\right) + \left(-\frac{3}{14} * \log_2 \left(\frac{3}{14}\right)\right) = 0.27336830983 + 0.047622694742142857142857142857143 = 0.74959525725142857142857142857143$$

$$\text{Entropy (2017)} = \left(-\frac{10}{13} * \log_2 \left(\frac{10}{13}\right)\right) + \left(-\frac{3}{13} * \log_2 \left(\frac{3}{13}\right)\right) = 0.29116278711538461538461538461538 + 0.48818705016923076923076923076923 = 0.77934983728461538461538461538461$$

### Entropy Jenis Kelamin

$$\text{Entropy (Laki-Laki)} = \left(-\frac{14}{17} * \log_2 \left(\frac{14}{17}\right)\right) + \left(-\frac{3}{17} * \log_2 \left(\frac{3}{17}\right)\right) = 0.23067710992117647058823529411765 + 0.44161770714705882352941176470588 = 0.67229481706823529411764705882353$$

$$\text{Entropy (PEREMPUAN)} = \left(-\frac{7}{10} * \log_2 \left(\frac{7}{10}\right)\right) + \left(-\frac{3}{10} * \log_2 \left(\frac{3}{10}\right)\right) = 0.360201220981 + 0.52108967826 = 0.881290899241$$

Entropy Nilai

$$\text{Entropy (Sedang)} = \left(-\frac{21}{27} * \log_2 \left(\frac{21}{27}\right)\right) + \left(-\frac{6}{27} * \log_2 \left(\frac{6}{27}\right)\right) = 0.28199895062888888888888888888889 + 0.48220555586666666666666666666667 = 0.76420450649555555555555555555556$$

Perhitungan Gain Akar

$$\text{Gain (Total, Angkatan)} = 1 - \left(\left(\frac{14}{27}\right) * 0.74959525725142857142857142857143\right) + \left(\frac{13}{27} * 0.77934983728461538461538461538461\right) =$$

$$1 - (0.38867902227851851851851851852 + 0.37524251424814814814814814814815) = 0.23607846347333333333333333333333$$

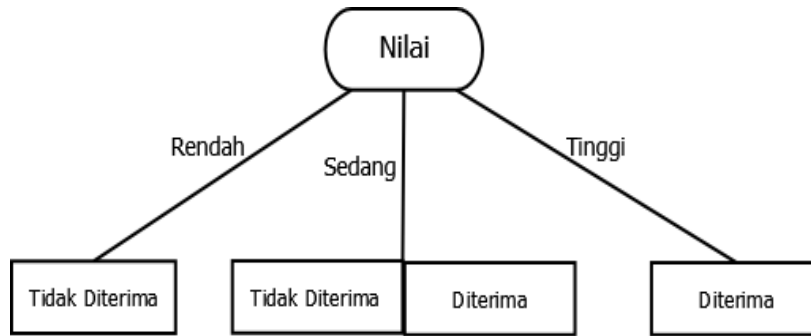
$$\text{Gain (Total, Jenis Kelamin)} = 1 - \left(\left(\frac{17}{27}\right) * 0.67229481706823529411764705882353\right) + \left(\frac{10}{27} * 0.881290899241\right) =$$

$$1 - (0.42329673667259259259259259259259 + 0.32640403675592592592592592592593) = 0.25029922657148148148148148148148$$

TABLE V. TABEL HASIL PERHITUNGAN ENTROPY DAN GAIN NODE 1.1

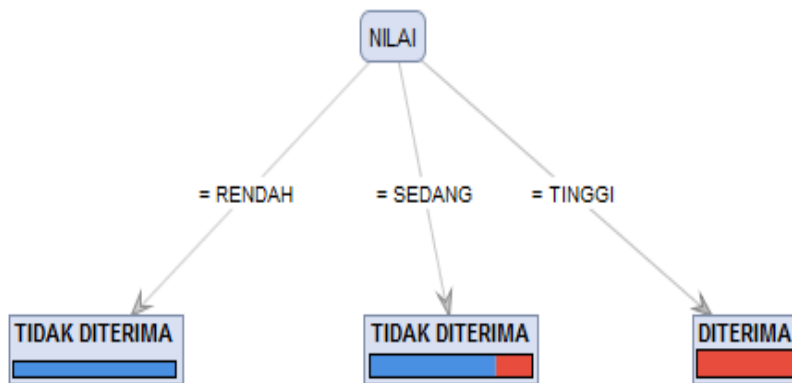
NODE	KRITERIA	ATRIBUT	JUMLAH KASUS	DITERIMA	TIDAK DITERIMA	ENTROPHY	GIAN
1.1	ANGKATAN						0.236078463 47333333333 33333333333 3
		2016	14	3	11	0.74959525 725142857 142857142 857143	
		2017	13	3	10	0.77934983 728461538 461538461 538461	
	JENIS KELAMIN						0.250299226 57148148148 14814814814 8
		LAKI-LAKI	17	3	14	0.67229481 706823529 411764705 882353	
		PEREMPUAN	10	3	7	0.88129089 9241	
	NILAI						
		SEDANG	27	6	21	0.76420450 649555555 555555555 555556	

Dari hasil tabel V pada node 1.1, dapat diketahui bahwa atribut Nilai yaitu cabang sedang mempunyai 2 pengklasifikasian yaitu ada yang Diterima dan ada yang Tidak diterima.



Gambar 2. Hasil pohon akar node 1.1

Gambar 2 dapat disimpulkan bahwa berdasarkan pohon keputusan dapat diketahui atribut Nilai sebagai root. Kemudian Nilai mempunyai cabang yaitu Rendah, sedang dan Tinggi. Disini nilai tinggi dan sebagian nilai sedang sangat berpengaruh untuk memprediksi kelulusan calon anggota HMIF



Gambar 3. Hasil pohon akar dari tools Rapid Miner v6.5.2

Gambar diatas merupakan hasil pengklasifikasian dari pohon akar, yang menjadi akar pohonnya atribut Nilai dan mempunya tiga cabang, yaitu rendah, sedang dan tinggi

```

PerformanceVector
PerformanceVector:
accuracy: 92.50%
ConfusionMatrix:
True:  TIDAK DITERIMA  DITERIMA
TIDAK DITERIMA: 40    6
DITERIMA:    0    34
precision: 100.00% (positive class: DITERIMA)
ConfusionMatrix:
True:  TIDAK DITERIMA  DITERIMA
TIDAK DITERIMA: 40    6
DITERIMA:    0    34
recall: 85.00% (positive class: DITERIMA)
ConfusionMatrix:
True:  TIDAK DITERIMA  DITERIMA
TIDAK DITERIMA: 40    6
DITERIMA:    0    34
AUC (optimistic): 1.000 (positive class: DITERIMA)
AUC: 0.961 (positive class: DITERIMA)
AUC (pessimistic): 0.921 (positive class: DITERIMA)
  
```

Gambar 4. Hasil PerformanceVectory dari tools Rapid Miner v6.5.2

Pada gambar 4 terdapat pendeskripsian hasil dari performance yang meliputi nilai accuracy, confusion matrix, precision recall, AUC (optimistic), AUC, AUC (pessimistic).

## KESIMPULAN

Berdasarkan pengujian model berbasis Algoritma Decision Tree C4.5 menggunakan framework Rapid Miner v.6.5.2. Didapat hasil penelitian dari 80 dataset yang diolah dan menggunakan 7 atribut membuktikan bahwa algoritma Decision Tree C4.5 memiliki kinerja yang lebih tinggi dibandingkan dengan algoritma lainnya. Hal ini karena secara keseluruhan algoritma C4.5 mempunyai nilai kinerja akurasi yang tinggi. Untuk penelitian selanjutnya dengan permasalahan yang sama, akan dilakukan penanganan terhadap pengklasifikasian atribut pada dataset yang mungkin dapat memberikan hasil yang lebih baik dalam keakuratan data.

Jadi dari hasil penelitian di atas akar dari pohon keputusan yaitu atribut Nilai kemudian mempunyai cabang dengan kategori nilai rendah (tidak diterima, nilai sedang sebagian (diterima) dan (tidak diterima), sedangkan nilai tinggi (diterima). Dan memiliki accuracy 92,50%, precision 100%, recall sebesar 85,00%, AUC (optimistic) sebesar 1000, AUC sebesar 0,961, sedangkan AUC (pessimistic) sebesar 0,921. Maka total penerimaan calon anggota HMIF yaitu, 40 diterima dan 40 tidak diterima. Jadi hasil dari Decision tree C4.5 menunjukkan bawah nilai tinggi sangat mempengaruhi kelulusan calon anggota.

## DAFTAR PUSTAKA

- [1] Larissa Navia Rani, 2016, *Klasifikasi Nasabah Menggunakan Algoritma C4.5 Sebagai Dasar Pemberian Kredit*, Jurnal Inovtek Polbeng - Seri Informatika, Vol. 1, No. 2, November 2016
- [2] Erlin Elisa, 2017, *Analisis dan Penerapan Algoritma C4.5 Dalam Data Mining Untuk Mengidentifikasi Faktor-Faktor Penyebab Kecelakaan Kerja Kontruksi PT.Arupadbatu Adisesanti*, JOIN, Volume 2 No. 1, hal 37
- [3] Rendragraha Kumara, Catur Supriyanto, 2015, *Klasifikasi Data Mining Untuk Penerimaan Seleksi Calon Pegawai Negeri Sipil 2014 Menggunakan Algoritma Decision Tree C4.5*, Tesis, Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro Semarang
- [4] Angga Raditya, *Implementasi Data Mining Classification untuk Mencari Pola Prediksi Hujan dengan Menggunakan Algoritma C4.5*
- [5] Ari Muzakir, Rika Anisa Wulandari, 2016, *Model Data Mining sebagai Prediksi Penyakit Hipertensi Kehamilan dengan Teknik Decision Tree*, *Scientific Journal of Informatics*, Vol. 3, No. 1, Mei 2016
- [6] Sigit Abdillah, 2011, *Penerapan Algoritma Decision Tree C4.5 Untuk Diagnosa Penyakit Stroke Dengan Klasifikasi Data Mining Pada Rumah Sakit Sanata Maria Pemalang*, Skripsi, Program Pasca Sarjana Ilmu Komputer Universitas Dian Nuswantoro, Semarang
- [7] [Rismayanti, 2018, *Decision Tree Penentuan Masa Studi Mahasiswa Prodi Teknik Informatika (Studi Kasus: Fakultas Teknik dan Komputer Universitas Harapan Medan)*, Jurnal Sistem Informasi, Volume: 02, Number: 01, April 2018 ISSN 2579-5341 (online)

- [8] M. Abdillah Luthfi, 2016, *Klasterisasi Daerah Rawan Kecelakaan Lalu Lintas Di Kota Semarang Menggunakan Chebysev Distance K-Means*, Skripsi, Program Pasca Sarjana Ilmu Komputer Universitas Dian Nuswantoro, Semarang