



Pemanfaatan Algoritma C4.5 untuk Mendukung Pemilihan Konsentrasi Studi yang Tepat di Teknik Informatika

Desyanti^{1,*}, Rudi Faisal²

¹Teknik Informatika, Sekolah Tinggi Teknologi Dumai, Riau, Indonesia

²Teknik Industri, Sekolah Tinggi Teknologi Dumai, Riau, Indonesia

Email: ^{1,*}desyanti734@gmail.com, ²faisalrudifaisal@gmail.com

Email Penulis Korespondensi: desyanti734@gmail.com

Abstrak—Pemilihan konsentrasi studi yang tepat pada Program Studi Teknik Informatika merupakan hal krusial dalam menunjang keberhasilan proses pembelajaran dan karier mahasiswa. Saat ini dalam pemilihan konstansi program studi teknik infomatika di Kampus X Kota Dumai masih membebaskan mahasiswa untuk memilih konstansi tanpa mempertimbangkan kemampuan yang dimiliki, akibatnya banyak mahasiswa yang merasa salah dalam memilih konstansi yang menyebabkan nilai mata kuliah menjadi turun. Penelitian ini menggunakan algoritma C4.5 untuk membangun pohon keputusan guna membantu mahasiswa menentukan konsentrasi studi berdasarkan variabel seperti jenis kelamin, jurusan SLTA, nilai mata kuliah, serta minat dan bakat. Metode ini diawali dengan proses *Knowledge Discovery in Databases* (KDD) yang mencakup seleksi data, pembersihan data, dan transformasi menggunakan skala Likert. Data dari 74 mahasiswa semester V dianalisis untuk menghasilkan aturan keputusan yang relevan. Hasil implementasi menunjukkan bahwa algoritma C4.5 mampu memberikan akurasi tinggi dalam menentukan konsentrasi studi yang sesuai. Sistem ini diharapkan dapat menjadi alat pendukung keputusan bagi mahasiswa dan institusi pendidikan dalam proses penjurusan. Dari hasil pengujian dapat diketahui nilai *precision*, *recall*, dan *accuracy* untuk setiap percobaan. Nilai *precision* tertinggi dari semua percobaan yaitu 80%, sedangkan nilai *precision* terendah dari semua percobaan yaitu 50%. Untuk Nilai *recall* dari semua percobaan yaitu 100%. Nilai *accuracy* tertinggi dari semua percobaan yaitu 95,45%, sedangkan nilai *accuracy* terendah dari semua percobaan yaitu 81,48%.

Kata Kunci: Algoritma C4.5; Data Mining; Konsentrasi Studi; Teknik Informatika; Pohon Keputusan.

Abstract—Choosing the right study concentration in the Informatics Engineering Study Program is crucial in supporting the success of the student's learning process and career. Currently, when choosing a contrast for the information engineering study program at Campus X This research uses the C4.5 algorithm to build a decision tree to help students determine their study concentration based on variables such as gender, high school major, course grades, and interests and talents. This method begins with the Knowledge Discovery in Databases (KDD) process which includes data selection, data cleaning, and transformation using a Likert scale. Data from 74 fifth semester students were analyzed to produce relevant decision rules. The implementation results show that the C4.5 algorithm is able to provide high accuracy in determining the appropriate study concentration. This system is expected to be a decision support tool for students and educational institutions in the majoring process. From the test results, the precision, recall and accuracy values for each experiment can be seen. The highest precision value from all experiments is 80%, while the lowest precision value from all experiments is 50%. The recall value for all experiments is 100%. The highest accuracy value from all experiments was 95.45%, while the lowest accuracy value from all experiments was 81.48%.

Keywords: C4.5 Algorithm; Data Mining; Study Concentration; Informatics Engineering; Decision Tree.

1. PENDAHULUAN

Perkembangan teknologi sudah merambah kesemua bidang, baik itu pendidikan, kesehatan, pemerintahan, hukum dan lain sebagainya. Komputer saat ini memiliki peran yang sangat luar biasa pentingnya untuk membantu mempermudah pekerjaan manusia, bahkan sekarang komputer telah banyak menggantikan pekerjaan manusia menggunakan teknologi Artificial Intelligence (AI) [1][2]. Pendidikan Tinggi merupakan lanjutan dari pendidikan menengah yang harus ditempuh oleh seseorang jika ingin mencapai kejenjang berikutnya dalam pendidikan. Perguruan tinggi menjadi jembatan untuk seseorang siap menjadi profesional dalam bidang keahlian tertentu sesuai dengan pilihannya[3]. Untuk dapat memilih jurusan yang sesuai keinginan tentulah bukan faktor yang mudah, harus berdasarkan minat dan bakat yang dimiliki oleh seseorang itu sendiri, begitu juga dengan pemilihan konsentrasi pada suatu jurusan[4]. Proses penentuan jurusan dilakukan dengan tujuan untuk menempatkan mahasiswa sesuai dengan kemampuannya, namun seringkali dalam memilih konstansi mahasiswa lebih cenderung asal-asalan, dengan kata lain ikut-ikutan teman, bingung menentukan mau pilih konsentrasi apa, asal pilih saja dan faktor lainnya. Hal ini yang menyebabkan mahasiswa salah dalam memilih konsentrasi sehingga mempengaruhi proses belajarnya [5].

Konsentrasi Mata Kuliah dalam sebuah Program Studi merujuk pada pilihan peminatan atau fokus studi yang dirancang untuk mendalami bidang tertentu secara sistematis dan terarah. Peminatan ini memberikan kesempatan bagi mahasiswa untuk mengkhususkan diri dalam area spesifik, sehingga dapat mengembangkan keahlian yang lebih mendalam sesuai dengan minat dan kebutuhan karier mereka di masa depan. Melalui konsentrasi ini, mahasiswa tidak hanya mendapatkan pemahaman teoritis tetapi juga keterampilan praktis yang relevan dengan bidang keahlian yang dipilih [6]. Universitas X yang ada di Kota Dumai sudah meluluskan banyak mahasiswa yang berkualitas, salah satu program studinya adalah teknik Informatika. Sama dengan program studi umumnya yang memiliki konsentrasi di setiap jurusan. Program Studi Informatika juga menawarkan 3 (tiga) bidang konsentrasi studi yaitu *Soft Computing*, *Multimedia*, dan *Computer Network*. Mata kuliah konsentrasi merupakan mata kuliah yang dapat menunjukkan bakat dari mahasiswa [7]. Namun pada penerapannya mahasiswa masih kebingungan dalam menentukan konsentrasi apa yang sesuai dengan kemampuannya. Selama ini proses penentuan konsentrasi mahasiswa ditentukan oleh kemauan mahasiswa itu sendiri bukan berdasarkan kemampuan mahasiswa. Hal ini yang menyebabkan mahasiswa terkadang salah memilih konsentrasi sehingga pada akhirnya bingung dalam menentukan proyek untuk tugas akhirnya. Untuk itu dibutuhkan suatu analisa yang

mampu membantu mahasiswa dalam memilih konsentrasi yang sesuai dengan kemampuannya berdasarkan kriteria yang ditentukan. Data mining adalah proses eksplorasi dan analisis mendalam terhadap kumpulan data besar untuk mengidentifikasi pola, hubungan, atau informasi tersembunyi yang relevan, dengan tujuan mendukung pengambilan keputusan yang lebih efektif melalui pemanfaatan teknik komputasi dan statistik canggih [8]. Data mining dapat diartikan sebagai proses menggali informasi baru yang bernilai dari kumpulan data dalam jumlah besar untuk mendukung pengambilan keputusan.

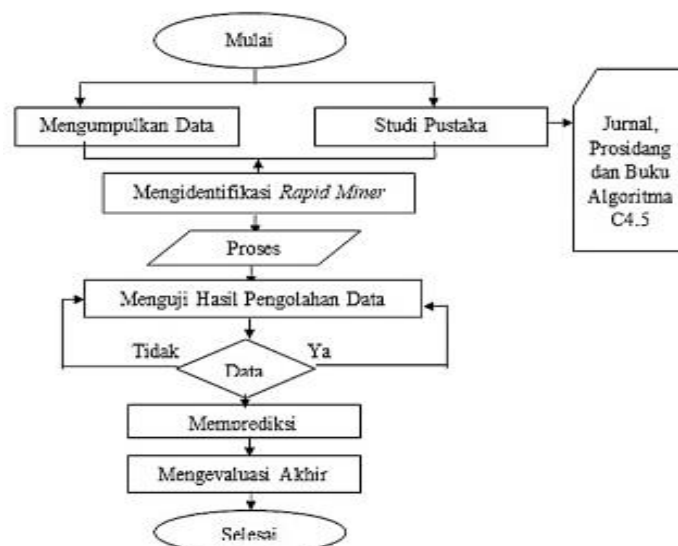
Data mining merupakan salah satu komponen utama dalam proses *Knowledge Discovery in Database (KDD)*, yang melibatkan beberapa tahap, seperti seleksi data, pra-pengolahan, transformasi, penerapan teknik data mining, dan evaluasi hasil. Dengan menerapkan data mining, wawasan penting yang tersembunyi dalam data besar dapat ditemukan dan dimanfaatkan secara optimal [9]. Salah satu metode data mining yang sesuai untuk mengatasi permasalahan pemilihan konsentrasi jurusan adalah algoritma C.45. Algoritma C4.5 adalah sebuah metode yang digunakan untuk membangun pohon keputusan. Pohon keputusan sendiri merupakan salah satu teknik yang populer dan efektif dalam proses klasifikasi serta prediksi. Pendekatan ini bekerja dengan mengorganisasi data dalam jumlah besar menjadi sebuah struktur pohon yang merepresentasikan aturan-aturan yang jelas dan mudah dipahami [10].

Penelitian sebelumnya dilakukan oleh ponti hariantio [11], ponti membahas tentang penentuan peminatan program studi informatika menggunakan variabel nilai mata kuliah, penalaran dan logika, sedangkan pada penelitian ini penulis akan menggunakan variabel Jenis Kelamin, Jurusan SLTA, dan Mata Kuliah Konsentrasi. Algoritma C.45 juga digunakan dalam menganalisa tingkat kecemasan mahasiswa dalam menyusun tugas akhir [12], hasil dari penelitian ini berupa tingkat kecemasan rendah, sedang dan tinggi yang dialami oleh mahasiswa dalam menyusun tugas akhir.

2. METODOLOGI PENELITIAN

2.1 Metode Pengembangan Sistem

Adapun tahapan penelitian yang akan dibuat dalam penelitian yang akan dilakukan terlihat pada Gambar 1.



Gambar 1. Kerangka Penelitian

a. Pengumpulan Data

Pada tahapan ini data diperoleh dari penyebaran kuisioner yang dilakukan oleh penulis dengan objek sasaran adalah mahasiswa semester V Program Studi Teknik Informatika dengan jumlah responden 74 mahasiswa. Hasil dari kuisioner diolah menggunakan skala likert. Skala Likert digunakan untuk mengukur jarak interval antara variabel satu dengan yang lain [19].

b. Studi Pustaka

Tahap ini mencari literatur yang berhubungan dengan penelitian yang dilakukan, studi pustaka dilakukan dengan mencari sumber – sumber referensi baik melalui buku dan jurnal.

c. Proses

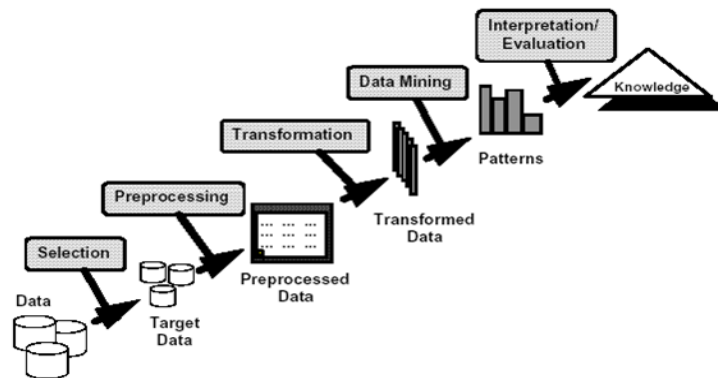
Melakukan perhitungan data kuisioner menggunakan metode C4.5 sesuai dengan bobot yang sudah di tentukan menggunakan Skala Likert, adapun yang menjadi variable ada penelitian ini adalah Jenis Kelamin, Jurusan SLTA, Pilihan Mata Kuliah Konsentrasi, Minat dan Bakat.

d. Menguji Data

Hasil dari perhitungan tersebut kemudian di uji menggunakan Rapid Miner

2.2 Data Mining

Dalam Data Mining ada beberapa tahapan yang merupakan proses dari *Knowledge Discovery in Databases* [KDD] seperti yang terlihat pada Gambar 2 [13].



Gambar 2. Proses *Knowledge Discovery in Databases* (KDD)

Dimana kegiatan ini dimulai dari menyeleksi data, kemudian memproses data, lalu mentransformasikan data yang selanjutnya akan diolah menggunakan metode data mining, hasil data yang diolah tersebut menghasilkan informasi baru yang dapat dijadikan acuan dalam penyelesaian masalah.

2.3 Algoritma C4.5

Algoritma C4.5 adalah sebuah metode yang digunakan untuk membangun pohon keputusan. Pohon keputusan sendiri merupakan teknik yang populer dan sangat efektif untuk proses klasifikasi dan prediksi. Metode ini bekerja dengan mengolah data dalam jumlah besar menjadi struktur pohon yang mencerminkan aturan-aturan dalam data tersebut [14]. Adapun langkah-langkah *algoritma C4.5* adalah sebagai berikut :

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

Dalam algoritma C4.5, Gain digunakan untuk menentukan atribut yang paling optimal dalam memisahkan objek. Sebelum menghitung Gain, langkah pertama adalah menentukan nilai Entropy. Entropy berfungsi untuk mengukur seberapa besar informasi yang dapat diberikan oleh suatu atribut masukan dalam memprediksi atribut keluaran. Secara sederhana, Entropy mencerminkan distribusi probabilitas dalam teori informasi, yang kemudian diadaptasi ke algoritma C4.5 untuk mengevaluasi tingkat homogenitas distribusi kelas dalam sebuah kumpulan data. Sebagai contoh, semakin tinggi nilai Entropy pada data, semakin seragam distribusi kelas di dalam kumpulan data tersebut. [15]

Sementara itu, perhitungan nilai entropi dapat dilihat pada persamaan 1. [12], [16]

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

Keterangan:

- S : himpunan kasus
A : fitur
n : jumlah partisi S
 p_i : proporsi dari S_i terhadap S

Setelah data set dibagi menjadi subset berdasarkan suatu atribut, nilai entropinya akan mengalami perubahan. Perubahan nilai entropi ini digunakan untuk mengevaluasi kualitas pembagian data yang telah dilakukan. Dalam algoritma C4.5, perubahan tersebut dikenal dengan istilah Gain.

Gain dihitung dengan mencari selisih antara entropi data set sebelum dan setelah pembagian (splitting) dilakukan. Pembagian yang optimal adalah yang menghasilkan entropi pada subset paling rendah, sehingga menghasilkan nilai Gain yang paling tinggi [17]. Pemilihan atribut sebagai akar dilakukan berdasarkan atribut yang memiliki nilai Gain tertinggi di antara semua atribut yang tersedia [18]. Untuk menghitung *Gain* digunakan rumus

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

Keterangan:

- S : himpunan kasus
A : atribut
n : jumlah partisi atribut A
 $|S_i|$: jumlah kasus pada partisi ke-i
 $|S|$: jumlah kasus dalam S



3. HASIL DAN PEMBAHASAN

3.1 Tahap Data Mining dengan Algoritma C4.5

Pada tahapan ini merupakan Gambaran proses analisa suatu masalah dan Gambaran dari penerapan metode untuk memecahkan masalah yang dihadapi. Untuk menunjang analisis data dalam pencarian *knowledge*, penulis akan menstransformasi data kuisioner yang sudah di isi oleh mahasiswa ke dalam bentuk *Microsoft Excel*. Berdasarkan penelitian penulis, penerapan *algoritma C4.5* yang digunakan untuk mengetahui kesesuaian pemilihan konsentrasi studi mahasiswa semester V terdiri dari variabel gender, jurusan SLTA, nilai mata kuliah, mata kuliah konsentrasi pilihan dan minat/bakat mahasiswa.

3.1.1 Selection Data

Pada penelitian ini populasi yang diteliti adalah Mahasiswa Informatika Semester V dengan jumlah reswponden 74 responden. Data yang diambil adalah jenis kelamin, jurusan SLTA, nilai mata kuliah yaitu matakuliah bersyarat dengan konsentrasi *Soft Computing*, Multimedia, *Computer Network*, mata kuliah konsentrasi pilihan dan minat/bakat mahasiswa. Data penelitian yang digunakan data responden mahasiswa semester V seperti terlihat pada Tabel 1.

Tabel 1. Data Responden Mahasiswa Semester V

No	JK	Jurusan SLTA	Pilihan Mata Kuliah Konsentrasi	Kom Graf	Multimedia		Soft comp	Computer Network		minat/b akat Soft Computing	Minat/ba kat Multimed ia	minat/ bakat Jaringan
					Mulmed	Prak Mulmed	KCB	Jarkom	Prak Jarkom			
1	L	Ipa	Multimedia	A	B	A	B	A	A	22	27	20
2	P	Ipa	Soft Computing	B	B	B	A	B	B	23	23	18
3	P	Inf	Multimedia	B	B	B	B	B	B	22	20	16
4	P	Ipa	Soft Computing	A	B	A	B	A	A	24	29	24
5	L	Ips	Soft Computing	A	A	A	A	A	A	20	24	17
6	L	Ipa	Multimedia							22	21	23
7	L	Ipa	Multimedia	A	A	A	B	B	B	19	22	10
8	L	TKJ	Multimedia							21	25	23
9	P	Marktg	Multimedia	A	A	A	A	A	A	24	26	21
10	P	Inf	Soft Computing	A	A	A	B	B	B	20	18	14
11	P	Ips	Soft Computing	A	A	A	B	A	A	21	22	17
74	P	Ips	Jaringan	B	B	C	B	B	C	22	21	22

Setelah didapat data responden , maka langkah selanjunya dilakukan pembersihan data yang dapat dilihat pada Tabel 2.

Tabel 2. Hasil Pembersihan Data

No	JK	Jurusan SLTA	Pilihan Mata Kuliah Konsentrasi	Kom Graf	Multimedia		Soft comp	Computer Network		minat/bakat Soft Computing	Minat/bakat Multimedia	minat/bakat jaringan
					Mulmed	Prak Mulmed	KCB	Jarkom	Prak Jarkom			
1	L	Ipa	Multimedia	A	B	A	B	A	A	22	27	20
2	P	Ipa	Soft Computing	B	B	B	A	B	B	23	23	18
3	P	Inf	Multimedia	B	B	B	B	B	B	22	20	16
4	P	Ipa	Soft Computing	A	B	A	B	A	A	24	29	24
5	L	Ips	Soft Computing	A	A	A	A	A	A	20	24	17
7	L	Ipa	Multimedia	A	A	A	B	B	B	19	22	10
8	P	Marktg	Multimedia	A	A	A	A	A	A	24	26	21
9	P	Inf	Soft Computing	A	A	A	B	B	B	20	18	14
10	P	Ips	Soft Computing	A	A	A	B	A	A	21	22	17
74	P	Ips	Jaringan	B	B	C	B	B	C	22	21	22

Tabel 2 menunjukan data yang lengkap, berisi tentang informasi mahasiswa yang akan memilih konsentrasi dan dapat diolah ke proses selanjutnya. Setelah dilakukan pembersihan data, selanjunya transformasi data menggunakan skala likert, adapun transformasi data dapat dilihat pada Tabel 3.

Tabel 3. Transformasi Data

No	JK	Jurusan SLTA	Pilihan Mata Kuliah Konsentrasi	Kom Graf	Multimedia		Soft comp	Computer Network		minat/bakat Soft Computing	Minat/bakat Multimedia	minat/bakat Jaringan
					Mulmed	Prak Mulmed	KCB	Jarkom	Prak Jarkom			
1	L	Ipa	Multimedia	A	B	A	B	A	A	Tinggi	Sangat Tinggi	Rendah
2	P	Ipa	Soft Computing	B	B	B	A	B	B	Tinggi	Tinggi	Rendah
3	P	Inf	Multimedia	B	B	B	B	B	B	Tinggi	Rendah	Rendah
4	P	Ipa	Soft Computing	A	B	A	B	A	A	Tinggi	Sangat Tinggi	Tinggi
5	L	Ips	Soft Computing	A	A	A	A	A	A	Rendah	Tinggi	Rendah



7	L	Ipa	Multimedia	A	A	A	B	B	B	Rendah	Tinggi	Sangat Rendah
8	P	Marktg	Multimedia	A	A	A	A	A	A	Tinggi	Tinggi	Tinggi
9	P	Inf	Soft Computing	A	A	A	B	B	B	Rendah	Rendah	Sangat Rendah
10	P	Ips	Soft Computing	A	A	A	B	A	A	Tinggi	Tinggi	Rendah
74	P	Ips	Jaringan	B	B	C	B	B	C	Tinggi	Tinggi	Tinggi

3.1.2 Implemetasi Algoritma C4.5

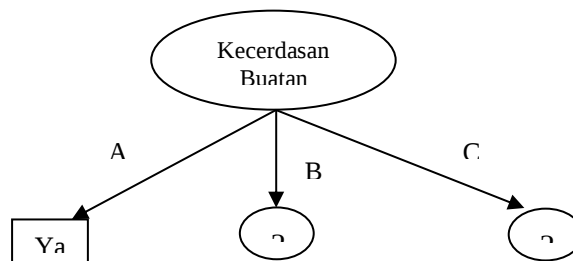
Data hasil Transfomasi dalam Tabel 3 digunakan untuk menghitung *Entropy* Total, Menghitung *Entropy* untuk atribut jenis kelamin dan Menghitung *Gain* untuk nilai jenis kelamin . Demikian juga untuk atribut lainnya, dilakukan proses perhitungan serupa, sehingga pada fase pencarian akar diperoleh dari hasil perhitungan seperti diperlihatkan dalam Tabel 4.

Tabel 4. Perhitungan Node 1

Node 1		Jml Kasus (S)	Tidak	Ya	<i>Entropy</i>	<i>Gain</i>
	TOTAL	45	9	36	0,721928095	
	Jenis Kelamin					0,006974316
	Laki-laki	14	2	12	0,591672779	
	Perempuan	31	7	24	0,770629069	
	Jurusan SLTA					0,149063123
	Ipa	13	3	10	0,779349837	
	Ips	11	4	7	0,945660305	
	Informatika	9	0	9	0	
	Teknik Komputer dan Jaringan	2	0	2	0	
	Rekayasa Perangkat Lunak	1	0	1	0	
	Marketing	2	0	2	0	
	Akuntansi	4	1	3	0,811278124	
	Adm Perkantoran	2	1	1	1	
	Teknik Otomotif R4	1	0	1	0	
	Mata Kuliah Konsentrasi					0,051141682
	Soft Computing	30	8	22	0,836640742	
	Multimedia	13	1	12	0,391243564	
	Computer Network	2	0	2	0	
	Komputer Grafis					0,182992299
	A	32	7	25	0,757878463	
	B	11	0	11	0	
	C	2	2	0	0	
	Multimedia					0,005715572
	A	24	4	20	0,650022422	
	B	21	5	16	0,791858353	
	C	0	0	0	0	
	Prak Multimedia					0,02591726
	A	27	4	23	0,605186577	
	B	16	4	12	0,811278124	
	C	2	1	1	1	
	Kecerdasan Buatan					0,288579814
	A	15	0	15	0	
	B	24	4	20	0,650022422	
	C	6	5	1	0,650022422	
	Jaringan komputer					0,027016745
	A	21	5	16	0,791858353	
	B	22	3	19	0,574635698	
	C	2	1	1	1	
	Prak Jarkom					0,052466679
	A	22	6	16	0,845350937	
	B	21	2	19	0,453716339	
	C	2	1	1	1	
	Minat/bakat softcomputing					0,057702805
	Sangat Tinggi	0	0	0	0	

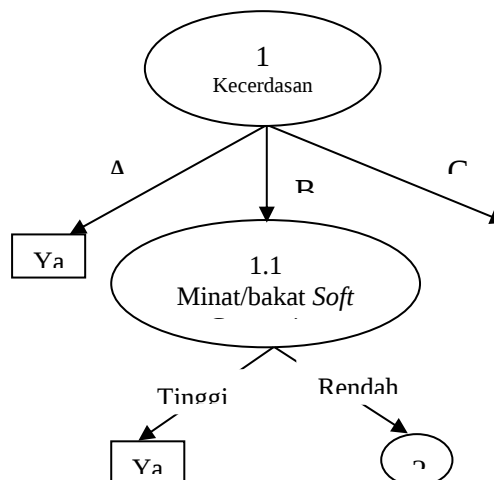
Minat/bakat multimedia	Tinggi	32	4	28	0,543564443
	Rendah	13	5	8	0,961236605
	Sangat Rendah	0	0	0	0
					0,081125574
Minat/bakat jaringan Komputer	Sangat Tinggi	4	0	4	0
	Tinggi	27	4	23	0,605186577
	Rendah	13	5	8	0,961236605
	Sangat Rendah	1	0	1	0
					0,089635369
	Sangat Tinggi	2	0	2	0
	Tinggi	15	1	14	0,353359335
	Rendah	26	8	18	0,89049164
	Sangat Rendah	2	0	2	0

Sesuai dengan hasil pada Tabel 4. dapat diketahui bahwa atribut dengan *Gain* tertinggi adalah Kecerdasan Buatan, yaitu sebesar 0,288579814. Dengan demikian, Kecerdasan Buatan dapat menjadi node akar. Ada tiga nilai atribut dari Kecerdasan Buatan, yaitu A, B, dan C. Dari ketiga nilai atribut tersebut, nilai atribut A sudah mengklasifikasikan kasus menjadi 1, yaitu keputusannya Ya dan untuk nilai atribut B dan C masih perlu dilakukan perhitungan lagi. Dari hasil tersebut dapat digambarkan pohon keputusan sementara tampak seperti Gambar 4.



Gambar 3. Pohon Keputusan Hasil Perhitungan *Entropy* Dan *Gain* Mencari Akar

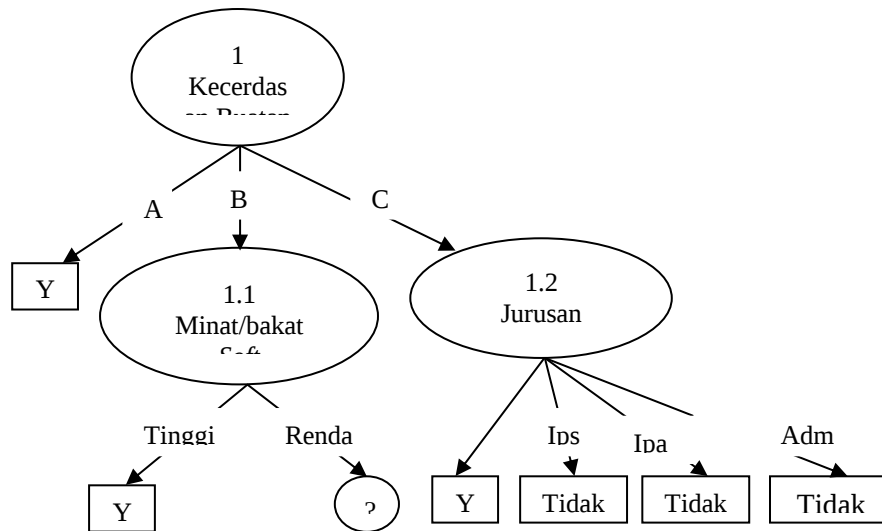
Selanjutnya akan mencari anak cabang dari nilai B, dari hasil perhitungan atribut dengan *Gain* tertinggi adalah Minat/bakat *Soft Computing* yaitu sebesar 0,245459674. Ada empat nilai atribut dari Minat/bakat *Soft Computing*, yaitu Sangat Tinggi, Tinggi, Rendah, dan Sangat Rendah. Dari empat nilai atribut tersebut, nilai atribut Tinggi sudah mengklasifikasikan kasus menjadi 1, yaitu keputusannya Ya dan nilai atribut Rendah masih belum mengklasifikasikan kasus dan untuk nilai atribut Sangat Tinggi dan Sangat Rendah tidak memiliki klasifikasi, sehingga tidak perlu dilakukan perhitungan lebih lanjut untuk atribut ini. Namun untuk nilai atribut Rendah masih dilakukan perhitungan lebih lanjut.



Gambar 4. Pohon Keputusan Hasil Perhitungan Node 1.1

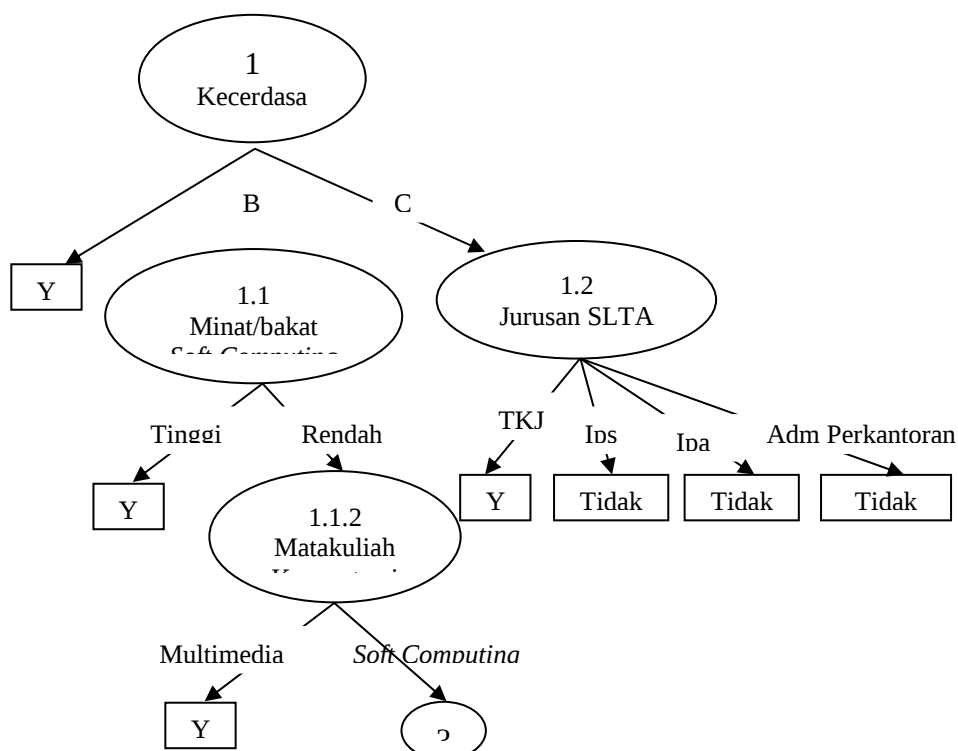
Pada tahap selanjutnya adalah untuk mencari anak cabang dari nilai C, dapat diketahui bahwa atribut dengan *Gain* tertinggi adalah Jurusan SLTA, yaitu sebesar 0,650022422. Dengan demikian, Jurusan SLTA dapat menjadi node cabang

dari nilai atribut C. Ada sembilan nilai atribut dari Jurusan SLTA, yaitu Ipa, Ips, Informatika, Teknik Komputer dan Jaringan, Rekayasa Perangkat Lunak, Marketing, Akuntansi, Adm Perkantoran, dan Teknik Otomotif R4. Dari kesembilan nilai atribut tersebut, nilai atribut Ipa, Ips dan Adm Perkantoran sudah mengklasifikasikan kasus menjadi 1, yaitu keputusannya Tidak, dan nilai atribut Teknik Komputer Jaringan sudah mengklasifikasikan kasus menjadi 1, yaitu keputusannya Ya dan untuk nilai atribut lainnya tidak memiliki klasifikasi, sehingga tidak perlu dilakukan perhitungan lebih lanjut.



Gambar 5. Pohon1 Keputusan Hasil Perhitungan Node 1.2

Pada tahap selanjutnya adalah untuk mencari anak cabang dari nilai Rendah diketahui bahwa atribut dengan *Gain* tertinggi adalah Matakuliah Konsentrasi Studi yaitu sebesar 0,609986547. Ada tiga nilai atribut dari matakuliah konsentrasi studi, yaitu *Softcomputing*, *Multimedia*, dan *Computer Network*. Dari tiga nilai atribut tersebut, nilai atribut *Multimedia* sudah mengklasifikasikan kasus menjadi 1, yaitu keputusannya Ya dan nilai atribut *Softcomputing* masih belum mengklasifikasikan kasus dan untuk nilai atribut *Computer Network* tidak memiliki klasifikasi, sehingga tidak perlu dilakukan perhitungan lebih lanjut untuk atribut ini. Namun untuk nilai atribut *Softcomputing* masih dilakukan perhitungan lebih lanjut



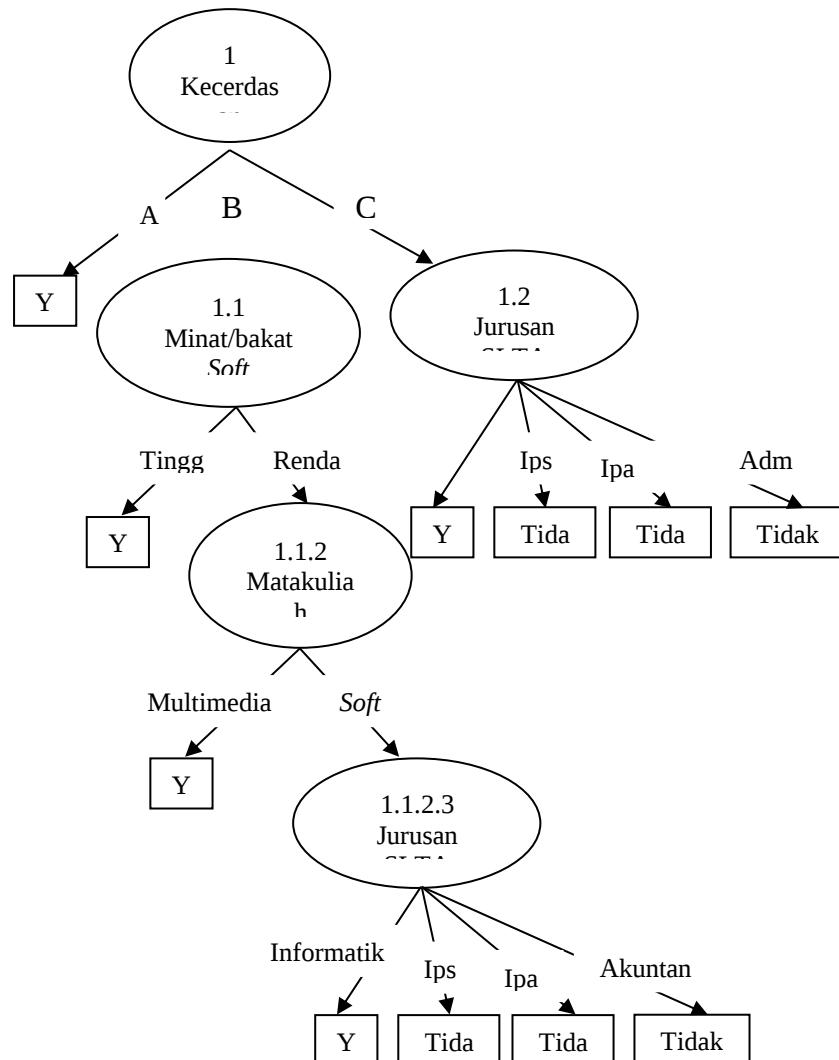
Gambar 6. Pohon Keputusan Hasil Perhitungan Node 1.1.2



Pada tahap selanjutnya adalah untuk mencari anak cabang dari nilai *Soft Computing* dapat diketahui bahwa atribut dengan *Gain* tertinggi adalah Jurusan SLTA, yaitu sebesar 0,721928095. Dengan demikian, Jurusan SLTA dapat menjadi node cabang dari nilai atribut C. Ada sembilan nilai atribut dari Jurusan SLTA, yaitu Ipa, Ips, Informatika, Teknik Komputer dan Jaringan, Rekayasa Perangkat Lunak, Marketing, Akuntansi, Adm Perkantoran, dan Teknik Otomotif R4. Dari kesembilan nilai atribut tersebut, nilai atribut Ipa, Ips dan Akuntansi sudah mengklasifikasikan kasus menjadi 1, yaitu keputusannya Tidak, dan nilai atribut Informatika sudah mengklasifikasikan kasus menjadi 1, yaitu keputusannya Ya dan untuk nilai atribut lainnya tidak memiliki klasifikasi, sehingga perlu dilakukan perhitungan lebih lanjut dan dapat dilihat pada gambar 7. Dari hasil pohon keputusan maka di dapatkan rule dari proses algoritma C4.5 sebagai berikut :

- Jika Kecerdasan Buatan = A maka status kesesuaian = Ya;
- Jika Kecerdasan Buatan = B, dan Minat/bakat *Soft Computing* = Tinggi maka status kesesuaian = Ya;
- Jika Kecerdasan Buatan = B, dan Minat/bakat *Soft Computing* = Rendah, dan Matakuliah konsentrasi = Multimedia maka status kesesuaian = Ya;
- Jika Kecerdasan Buatan = B, dan Minat/bakat *Soft Computing* = Rendah, dan Matakuliah konsentrasi = *Soft Computing*, dan Jurusan SLTA = Informatika maka status kesesuaian = Ya;
- Jika Kecerdasan Buatan = B, dan Minat/bakat *Soft Computing* = Rendah, dan Matakuliah konsentrasi = *Soft Computing*, dan Jurusan SLTA = Ips maka status kesesuaian = Tidak;
- Jika Kecerdasan Buatan = B, dan Minat/bakat *Soft Computing* = Rendah, dan Matakuliah konsentrasi = *Soft Computing*, dan Jurusan SLTA = Ipa maka status kesesuaian = Tidak;
- Jika Kecerdasan Buatan = B, dan Minat/bakat *Soft Computing* = Rendah, dan Matakuliah konsentrasi = *Soft Computing*, dan Jurusan SLTA = Akuntansi maka status kesesuaian = Tidak;
- Jika Kecerdasan Buatan = C dan Jurusan SLTA = TKJ (Teknik Komputer dan Jaringan) maka status kesesuaian = Ya;
- Jika Kecerdasan Buatan = C dan Jurusan SLTA = Ips maka status kesesuaian = Tidak;
- Jika Kecerdasan Buatan = C dan Jurusan SLTA = Ipa maka status kesesuaian = Tidak;
- Jika Kecerdasan Buatan = C dan Jurusan SLTA = Adm Perkantoran maka status kesesuaian = Tidak;

Berikut merupakan hasil pohon terakhir yang dihitung



Gambar 7. Pohon Keputusan Hasil Perhitungan Node 1.1.2.3



4. KESIMPULAN

Penelitian ini membuktikan bahwa algoritma C4.5 merupakan metode yang efektif untuk mendukung pemilihan konsentrasi studi mahasiswa Teknik Informatika. Dengan memanfaatkan data terkait minat, bakat, dan nilai akademik, algoritma ini menghasilkan pohon keputusan yang mampu mengklasifikasikan konsentrasi studi secara akurat. Hasil implementasi pada data mahasiswa semester V menunjukkan bahwa sistem ini dapat membantu mahasiswa memilih konsentrasi yang sesuai dengan kemampuan mereka, sehingga dapat mengurangi potensi kesalahan pemilihan yang sering terjadi. Namun, keterbatasan penelitian ini terletak pada jumlah data yang terbatas dan fokus pada variabel tertentu. Untuk penelitian selanjutnya, disarankan untuk memperluas variabel dan jumlah sampel guna meningkatkan keandalan sistem. Penerapan algoritma ini di lingkungan pendidikan diharapkan dapat meningkatkan efisiensi proses penentuan konsentrasi studi dan mendukung pengembangan karier mahasiswa secara optimal.

REFERENCES

- [1] S. M. Fifi Maqfiroh, "PENERAPAN KLASIFIKASI ALGORITMA DATA MINING C4.5 UNTUK MEMPREDIKSI TINGKAT KELULUSAN SISWA DI LEMBAGA PELATIHAN KERJA SHINJU SEMARANG," vol. 20, no. 1, pp. 105–123, 2022.
- [2] A. Winanda and S. Arfida, "Penerapan Algoritma Knuth Morris Pratt Terhadap Kelompok Bidang Keilmuan Mahasiswa Program Studi Teknik Informatika," *Pros. Semin. Nas. ...*, pp. 109–117, 2023, [Online]. Available: <https://jurnal.darmajaya.ac.id/index.php/PSND/article/view/3829%0Ahttps://jurnal.darmajaya.ac.id/index.php/PSND/article/viewFile/3829/1645>
- [3] I. Hidayanti, T. B. Kurniawan, and A. Afriyudi, "Perbandingan Dan Analisis Metode Klasifikasi Untuk Menentukan Konsentrasi Jurusan," *J. Ilm. Inform. Glob.*, vol. 11, no. 1, pp. 16–21, 2020, doi: 10.36982/jiig.v11i1.1067.
- [4] A. Baktiar, "Decission Tree Sebagai Metode Penentuan Penjurusan Perguruan Tinggi Berdasarkan Minat Dan Bakat Melalui Data Raport Dengan Uji Algoritma C4.5 (Studi Kasus di SMKN 1 Donorojo Pacitan)," *J. PILAR Teknol. J. Ilm. Ilmu Ilmu Tek.*, vol. 7, no. 1, pp. 40–45, 2022, doi: 10.33319/piltek.v7i1.110.
- [5] P. Subekti, T. D. Andini, and M. Islamiyah, "Sistem Penentuan Konsentrasi Jurusan Bagi Mahasiswa Informatika Menggunakan Metode K-Means Di Institut Asia Malang," *J. Manaj. Inform.*, vol. 12, no. 1, pp. 25–39, 2022, doi: 10.34010/jamika.v12i1.6452.
- [6] D. O. Sihombing, P. Studi, T. Informasi, I. Teknologi, and S. Setia, "Sistem Pendukung Keputusan Pemilihan Konsentrasi Mata Kuliah dengan Metode MOORA," vol. 5, no. 4, pp. 942–956, 2024, doi: 10.47065/josyc.v5i4.5780.
- [7] A. Ibnu Raharjo, N. Ramsari, and M. H. Rahman, "Sistem Pendukung Keputusan Untuk Pemilihan Konsentrasi Peminatan Menggunakan Metode Naïve Bayes (Studi Kasus Program Studi Teknik Informatika Universitas Nurtanio Bandung)," *J. Teknol. Inf. dan Komun.*, vol. 12, no. 2, 2023, doi: 10.56244/fiki.v12i2.654.
- [8] I. A. Gellysa Urva, Desyanti, *PENERAPAN DATA MINING DI BERBAGAI BIDANG : Konsep, Metode, dan Studi Kasus*. 2023. [Online]. Available: https://books.google.co.id/books/about/PENERAPAN_DATA_MINING_DI_BERBAGAI_BIDANG.html?id=uq6-EAAAQBAJ&redir_esc=y
- [9] C. Nas, "Data Mining Prediksi Minat Calon Mahasiswa Memilih Perguruan Tinggi Menggunakan Algoritma C4.5," *J. Manaj. Inform.*, vol. 11, no. 2, pp. 131–145, 2021, doi: 10.34010/jamika.v11i2.5506.
- [10] I. A. C., "UNTUK MEMPREDIKSI KESESUAIAN GAYA BELAJAR SISWA SDN 02 PESIDO BERBASIS JAVA," vol. 05, no. 04, pp. 742–749, 2024.
- [11] P. Harianto, "Penerapan Metode Fuzzy Mamdani Untuk Penentuan Peminatan Program Studi Teknik Informatika di STMIK Pontianak," *Semin. Nasional Corisindo*, pp. 134–138, 2023.
- [12] D. Desyanti, "Implementasi Metode C.45 dalam Menganalisa Tingkat Kecemasan Mahasiswa Menyusun Tugas Akhir," *J. Unitek*, vol. 14, no. 1, pp. 17–29, 2021, doi: 10.52072/unitek.v14i1.175.
- [13] M. Kamil and W. Cholil, "Analisis Perbandingan Algoritma C4.5 dan Naive Bayes pada Lulusan Tepat Waktu Mahasiswa di Universitas Islam Negeri Raden Fatah Palembang," *J. Inform.*, vol. 7, no. 2, pp. 97–106, 2020, doi: 10.31294/ji.v7i2.7723.
- [14] D. Desyanti, "SATIN-Sains dan Teknologi Informasi Penerapan Data Mining Algoritma C4.5 untuk Mengetahui Tingkat Kepuasan Konsumen di Hotel Grand Zuri Dumai," *SATIN - Sains dan Teknol. Inf.*, vol. 4, no. 2, 2018, [Online]. Available: <http://jurnal.stmik-amik-riau.ac.id>
- [15] I. M. Rahmawati and H. Wibowo, "Penerapan Algoritma C4.5 Dalam Memprediksi Peminatan Mahasiswa Program Studi Teknik Informatika," *Ijccs*, vol. x, No.x, no. x, pp. 1–5, 2023, [Online]. Available: <https://jurnal.polsri.ac.id/index.php/teknika/article/view/7355%0Ahttps://jurnal.polsri.ac.id/index.php/teknika/article/view/7355/2812>
- [16] F. Annisa Saburina, Husaini, "SISTEM PENDUKUNG KEPUTUSAN PEMILIHAN MATA KULIAH MBKM DI UNIVERSITAS JABAL GHAFUR MENGGUNAKAN ALGORITMA C4.5," no. 1, pp. 10–17.
- [17] S. Mawaddah, W. J. Pranoto, and F. Faldi, "Optimasi Algoritma C4.5 Menggunakan Metode Adaboost Classification Pada Klasifikasi Nilai Mahasiswa Studi Kasus: Universitas Muhammadiyah Kalimantan Timur," *J. Sains Komput. dan Teknol. Inf.*, vol. 6, no. 1, pp. 83–89, 2023, doi: 10.33084/jsakti.v6i1.5458.
- [18] T. Widiastuti, K. Karsa, and C. Juliane, "Evaluasi Tingkat Kepuasan Mahasiswa Terhadap Pelayanan Akademik Menggunakan Metode Klasifikasi Algoritma C4.5," *Technomedia J.*, vol. 7, no. 3, pp. 364–380, 2022, doi: 10.33050/tmj.v7i3.1932.
- [19] D. Desyanti, Y. Yusrizal, and F. Sari, "Implementasi Algoritma K-Modes Untuk Mengukur Tingkat Kepuasan Mahasiswa Terhadap Pembelajaran Daring," *Build. Informatics, Technol. Sci.*, vol. 3, no. 4, pp. 719–727, 2022, doi: 10.47065/bits.v3i4.1401.