

# ALGORITMA C4.5 UNTUK MENENTUKAN KLASIFIKASI TINGKAT PEMAHAMAN MAHASISWA PADA MATAKULIAH BAHASA PEMROGRAMAN

**Dhea Halimah**

STIKOM Tunas Bangsa, Pematangsiantar, Indonesia

Korespondensi penulis: [dheahalimah19@gmail.com](mailto:dheahalimah19@gmail.com)

**Muhammad Ridwan Lubis**

STIKOM Tunas Bangsa, Pematangsiantar, Indonesia

**Widodo Saputra**

STIKOM Tunas Bangsa, Pematangsiantar, Indonesia

**Abstract.** *The purpose of this research is to classify students in determining the level of understanding of students in programming language courses using the C4.5 datamining method. By knowing the level of student understanding, the study program can provide useful input as a consideration for lecturers in carrying out student teaching and learning programs and can further improve learning if the level of understanding is good. Sources of data obtained from the results of a questionnaire given to students of STIKOM Tunas Bangsa Pematangsiantar semester 4 and semester 6 of the Information Systems Study Program. The criteria used include Student Interest (C1), Motivation (C2), Communication (C3), Learning Media (C4), Facilities and Infrastructure (C5). The research test process uses RapidMiner software to create a decision tree. From the results of processing C4.5 using RapidMiner software, attribute C2 (Motivation) is the most influential attribute to improve students' understanding of programming language courses and the performance data shown against the suitability of the C4.5 method has an accuracy of 84.38%.*

**Keywords:** *C4.5 Algorithm, Data Mining, Classification, Courses, Level of Understanding.*

**Abstrak.** Penelitian ini bertujuan melakukan klasifikasi dalam menentukan tingkat pemahaman mahasiswa pada matakuliah bahasa pemrograman dengan menggunakan metode data mining C4.5. Dengan mengetahui tingkat pemahaman mahasiswa, pihak prodi dapat memberi masukan yang bermanfaat sebagai bahan pertimbangan kepada para dosen dalam melaksanakan program kegiatan belajar mengajar terhadap mahasiswa dan dapat lebih meningkatkan pembelajaran apabila tingkat pemahaman sudah baik. Sumber data diperoleh dari hasil kuesioner yang diberikan kepada Mahasiswa STIKOM Tunas Bangsa Pematangsiantar semester 4 dan semester 6 prodi sistem informasi. Adapun Kriteria yang digunakan diantaranya Minat Mahasiswa (C1), Motivasi (C2), Komunikasi (C3), Media Pembelajaran (C4), Sarana dan Prasarana (C5). Proses uji penelitian menggunakan software RapidMiner untuk membuat pohon keputusan. Dari hasil

*Received Juli 07, 2022; Revised Agustus 2, 2022; September 22, 2022*

\* Dhea Halimah, [dheahalimah19@gmail.com](mailto:dheahalimah19@gmail.com)

pengolahan C4.5 dengan menggunakan bantuan software RapidMiner adalah atribut C2 (Motivasi) menjadi atribut yang paling berpengaruh untuk meningkatkan pemahaman mahasiswa pada matakuliah bahasa pemrograman dan data performance yang ditunjukkan terhadap kesesuaian metode C4.5 akurasinya adalah 84.38%.

**Kata kunci:** Algoritma C4.5, Data Mining, Klasifikasi, Matakuliah, Tingkat Pemahaman.

## **LATAR BELAKANG**

Bahasa Pemrograman (*programming language*) yang sering disebut sebagai bahasa komputer yang merupakan suatu instruksi standar yang digunakan untuk memerintahkan komputer secara langsung untuk menjalankan sebuah program atau perintah khusus (Anton, 2017). Semua ilmu harus dipelajari dari dasar dan membutuhkan proses untuk hasil yang maksimal. Dalam jurusan komputer mata kuliah bahasa pemrograman merupakan salah satu mata kuliah yang wajib di pelajari. Melalui pembelajaran mata kuliah bahasa pemrograman mahasiswa diharapkan dapat menumbuhkan kemampuan pemahaman sehingga dapat menggunakan bahasa pemrograman untuk membantu menyelesaikan masalah yang kompleks.

Bahasa Pemrograman merupakan notasi yang digunakan untuk menulis program di komputer. Bahasa Pemrograman adalah salah satu matakuliah penting di Perguruan Tinggi berbasis komputer (Zer, Hartama, & Andani, 2019). Mata kuliah bahasa pemrograman diajarkan secara teori dan praktik. Seperti yang diketahui dalam bahasa pemrograman memiliki tingkat kesulitan yang berbeda, sehingga sering dianggap mahasiswa sebagai mata kuliah yang sulit untuk dipahami. Selain itu banyak mahasiswa yang menganggap bahwa bahasa pemrograman hanya bisa dipelajari oleh mahasiswa yang pintar. Mahasiswa merupakan salah satu faktor utama dalam proses belajar mengajar (Raharjo & Windarto, 2021). Tidak hanya mahasiswa yang memiliki kemauan belajar yang tinggi yang menjadi peran pentingnya, dosen juga memiliki peran penting yang utama dalam menyampaikan materi perkuliahan agar dapat di pahami oleh mahasiswa. Terutama yang berkaitan bagaimana cara seorang dosen itu menyampaikan isi dari materi perkuliahan. Setiap dosen memiliki cara berbeda pula dalam penyampaian materi kepada mahasiswa mereka, ketika proses perkuliahan berlangsung penyampaian materi yang diberikan yang sangat berpengaruh pada hasil yang akan diperoleh mahasiswa.

Banyaknya konsep-konsep yang harus dipahami membuat mahasiswa kesulitan dalam menyelesaikan persoalan bahasa pemrograman (Novika, Poningsih, Okprana, Windarto, & Siahaan, 2021). Seperti halnya beberapa mahasiswa jurusan sistem informasi STIKOM Tunas Bangsa, mereka masih merasa kesulitan memahami matakuliah bahasa pemrograman. Seperti proses menulis kode-kode program menggunakan bahasa pemrograman tertentu, melakukan tahapan rilis dalam pengembangan suatu program, melakukan debugging untuk mengidentifikasi kesalahan yang ada di dalam programnya, bila ada kesalahan yang ditemukan perlu melakukan perbaikan dan merilis versi program yang baru, melakukan proteksi secara berkala agar tidak ada celah kesalahan di dalam programnya dan melakukan update. Mahasiswa mudah paham atau sulit memahami mata kuliah bahasa pemrograman dikarenakan

beberapa faktor yang mempengaruhi mahasiswa untuk paham dengan matakuliah bahasa pemrograman. Beberapa faktor yang mempengaruhi pemahaman konsep dalam bahasa pemrograman yaitu faktor internal yang berasal dari mahasiswa itu sendiri seperti minat mahasiswa, motivasi mahasiswa dalam belajar, cara belajar mahasiswa dan faktor eksternal yaitu faktor yang berasal dari luar mahasiswa seperti sarana dan prasarana yang mendukung belajar mahasiswa, media pembelajaran yang digunakan dan penyampaian dosen dalam proses pembelajaran.

Untuk mengatasi permasalahan tersebut dilakukan sebuah penelitian dengan menggunakan *Data Mining* klasifikasi dengan Algoritma C4.5. untuk mendapatkan solusi yang lebih efisien pada tingkat pemahaman mahasiswa pada matakuliah bahasa pemrograman. Algoritma C4.5 merupakan salah satu algoritma yang digunakan untuk melakukan klasifikasi atau segmentasi yang bersifat prediktif (Sitorus, Windarto, Hartama, & Damanik, 2019). Dasar Algoritma C4.5 adalah pembentukan pohon keputusan (*Decision Tree*). Kelebihan dari metode ini dapat membentuk sebuah pohon keputusan yang mudah dipahami, dapat membuat klasifikasi lebih mudah, sederhana dan juga memiliki tingkat akurasi yang tinggi.

## KAJIAN TEORITIS

### 1. Data Mining

Data mining adalah proses yang mempekerjakan satu atau lebih teknik pembelajaran komputer (*machine learning*) untuk menganalisis dan mengekstraksi pengetahuan (*knowledge*) secara otomatis (Riandari & Simangunsong, 2019). Data mining disebut juga dengan *Knowledge Discovery in Database (KDD)*, yang merupakan kegiatan meliputi pengumpulan, pemakaian data dalam menemukan keteraturan dari suatu proses untuk mencari nilai tambahan dari sekumpulan data dengan pola dan hubungan dalam data yang berkapasitas besar.

Menurut Jiawei Han dalam buku *Data Mining Concepts and Techniques KDD* atau *Knowledge Discovery from Data*, merupakan proses terstruktur, yaitu sebagai berikut:

- a) *Data Cleaning* adalah Proses membersihkan data dari data *noise* dan tidak konsisten.
- b) *Data Integration* adalah Proses untuk menggabungkan data dari beberapa sumber yang berbeda.
- c) *Data Selection* adalah Proses untuk memilih data dari database yang sesuai dengan tujuan analisis.
- d) *Data Transformation* adalah Proses mengubah bentuk data menjadi data yang sesuai untuk proses Mining.
- e) *Data Mining* adalah Proses penting yang menggunakan sebuah metode tertentu untuk memperoleh sebuah pola dari data.
- f) *Pattern Evaluation* adalah Proses mengidentifikasi pola.
- g) *Knowledge Presentation* adalah yang dapat merepresentasikan informasi yang dibutuhkan, proses dimana informasi yang telah didapatkan kemudian digunakan oleh pemilik data (Sulastri & Gufroni, 2017).

## 2. Decision Tree (Pohon Keputusan)

Pohon keputusan merupakan salah satu metode klasifikasi yang paling populer karena mudah untuk diinterpretasi oleh manusia, model prediksi menggunakan struktur pohon atau struktur berhirarki. Konsep dari pohon keputusan adalah mengubah data menjadi pohon keputusan dan aturan aturan keputusan (Junia & Riandari, 2019).

*Decision tree* (pohon keputusan) adalah sebuah struktur yang dapat digunakan untuk membagi kumpulan data yang besar menjadi himpunan-himpunan record yang lebih kecil dengan menerapkan serangkaian aturan keputusan (Setio, Saputro, & Bowo Winarno, 2020). Hasil dari proses perhitungan entropy dan information gain, setelah perhitungan berulang-ulang sampai semua atribut pohon memiliki kelas dan tidak bisa lagi dilakukan proses perhitungan (Cynthia & Ismanto, 2018).

Ada 3 jenis node di *decision tree*, yaitu :

- a) *Root node*, yang merupakan *node* teratas, *node* ini tidak memiliki *input* dan tidak dapat memiliki *output* atau banyak *output*.
- b) *Internal node*, yang merupakan *node* cabang, *node* ini hanya memiliki satu *input* dan setidaknya dua *output*.
- c) *Leaf node* atau terminal *node*, ini merupakan *node* akhir, *node* ini hanya memiliki satu *input* dan tidak memiliki *output*.

## 3. Algoritma C4.5

Algoritma C4.5 merupakan kelompok algoritma *decision tree*. Algoritma ini mempunyai input berupa *training samples* dan *samples*. *Training samples* berupa data contoh yang akan digunakan untuk membangun sebuah *tree* yang telah diuji kebenarannya. Sedangkan *samples* merupakan *field-field* data yang nantinya akan kita gunakan sebagai parameter dalam melakukan klasifikasi data (Takalapeta, 2018).

Secara umum terdapat empat langkah dalam proses pembuatan pohon keputusan pada algoritma C4.5 yaitu :

- a) Memilih atribut sebagai akar (*rootnode*)

Langkah – langkah memperoleh atribut sebagai akar adalah dengan menghitung jumlah kasus dan jumlah target atribut. Kemudian menghitung nilai *entropy* yang digunakan untuk menghasilkan nilai atribut.

Berikut rumus menghitung *entropy* :

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

Dimana:

S : himpunan kasus

n : jumlah partisi S

$p_i$  : proporsi dari  $S_i$  terhadap S

setelah menghitung nilai *entropy*, selanjutnya menghitung *gain* untuk pemisahan objek.

Dengan menggunakan rumus berikut ini:

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

Keterangan:

S : Himpunan kasus

A : Atribut

n : Jumlah partisi atribut A

$|S_i|$  : Jumlah kasus pada partisi ke i

$|S|$  : Jumlah kasus dalam S

Setelah keseluruhan atribut dihitung menggunakan persamaan diatas, maka atribut yang memiliki *gain* tertinggi dibanding dengan atribut lain dijadikan sebagai *node* (akar).

b) Membuat cabang

Setelah diperoleh atribut yang memiliki *gain* tertinggi, maka atribut tersebut dijadikan *node*. *Node* ini memiliki *instance* sehingga *instance* dijadikan cabang dari *node*.

c) Membagi setiap kasus dalam cabang

Setiap nilai *instance* diklasifikasikan berdasarkan makna dari nilai *instance* agar menjadi lebih sederhana. Jika nilai *instance* tidak dapat disederhanakan lagi maka perlu melakukan perhitungan lebih lanjut.

d) Mengulangi proses

Ulangi proses untuk setiap cabang sampai semua cabang mempunyai kelas yang sama.

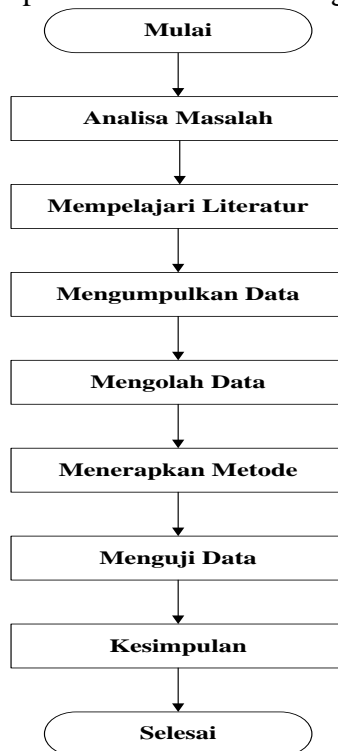
## METODE PENELITIAN

### 1. Lokasi dan Waktu Penelitian

Penelitian ini dilakukan di STIKOM Tunas Bangsa yang berlokasi di Jl. Jend. Sudirman Blok A No.1, 2 dan 3 Pematangsiantar. Waktu pengumpulan data dilakukan pada tanggal 20 Mei 2022 sampai, 26 Mei 2022 dengan memberikan pertanyaan melalui kuisioner kepada mahasiswa STIKOM Tunas Bangsa dengan mengambil sampel terhadap beberapa kelas sebagai perwakilan dari Program Studi Sistem Informasi.

### 2. Rancangan Penelitian

Rancangan penelitian dapat dilihat dalam rancangan *flowchart* pada gambar 1.



Gambar 1. Rancangan Penelitian

Berdasarkan rancangan penelitian pada gambar 3.1. diatas dapat disimpulkan Algoritma penelitian yang dilakukan adalah sebagai berikut :

- a) Analisa Masalah
- b) Mempelajari Literatur
- c) Mengumpulkan Data
- d) Mengelola Data
- e) Menerapkan Metode
- f) Menguji Data
- g) Kesimpulan

### **3. Prosudur Pengumpulan Data**

Pengumpulan data yang digunakan untuk penelitian terdapat beberapa metode yang terdiri dari :

- a) Penelitian Kepustakaan (*Library Research*) yaitu memanfaatkan perpustakaan, buku, prosiding atau jurnal sebagai media untuk referensi dalam menentukan faktor, parameter dan label yang digunakan dalam penelitian.
- b) Penelitian Lapangan (*Field Work Research*) yaitu penelitian yang dilakukan secara langsung ke lapangan dengan menggunakan beberapa teknik sebagai berikut :

- 1) Studi Literatur

Informasi yang penulis peroleh dengan mengumpulkan data, mempelajari data, validasi data dan mencari referensi terkait dengan kasus penelitian. Hasil dari studi literatur ini adalah tersusun dan terkoleksinya referensi yang baik dan benar dengan penelitian.

- 2) Kuisisioner

Kuisisioner yang dilakukan dengan pengumpulan data atau informasi kepada pihak yang mengisi jawaban dari pertanyaan yang diberikan kepada mahasiswa semester 4 dan semester 6 STIKOM Tunas Bangsa Pematangsiantar Prodi Sistem Informasi.

### **4. Analisis Data**

Pada penelitian ini penulis memperoleh data dari hasil kuisisioner yang diberikan kepada mahasiswa STIKOM tunas bangsa prodi sistem informasi semester 4 dan semester 6. Jumlah sampel data sebanyak 110 responden.

## **HASIL DAN PEMBAHASAN**

### **1. Pengolahan Data**

Perhitungan Algoritma C4.5 dimulai dengan menghitung jumlah kasus untuk keputusan paham atau tidak paham. Menghitung *entropy* dari semua kasus yang dibagi berdasarkan atribut Minat Mahasiswa (C1), Motivasi (C2), Komunikasi (C3), Media Pembelajaran (C4), Sarana Prasarana (C5). Kemudian menghitung *entropy* dengan persamaan (1) dan menghitung *gain* untuk masing-masing atribut dengan persamaan (2). Berikut ini adalah perhitungan nilai *entropy* dan *gain*.

**Menghitung *entropy* total :**

$$Entropy (Total) = \left( -\frac{(paham)}{total} * \log_2 \left( \frac{(paham)}{total} \right) \right) + \left( -\frac{(tidak)}{total} * \log_2 \left( \frac{(tidak)}{total} \right) \right)$$

$$\text{Entropy [Total]} = \left(-\frac{76}{110} x \log_2 \left(\frac{76}{110}\right)\right) + \left(-\frac{34}{110} x \log_2 \left(\frac{34}{110}\right)\right)$$

$$\text{Entropy [Total]} = 0,89212128$$

**Menghitung entropy dan gain Minat Mahasiswa :**

$$\text{Entropy [Ada]} = \left(-\frac{76}{94} x \log_2 \left(\frac{76}{94}\right)\right) + \left(-\frac{18}{94} x \log_2 \left(\frac{18}{94}\right)\right)$$

$$= 0,70457671$$

$$\text{Entropy [Tidak Ada]} = \left(-\frac{0}{16} x \log_2 \left(\frac{0}{16}\right)\right) + \left(-\frac{16}{16} x \log_2 \left(\frac{16}{16}\right)\right)$$

$$= 0$$

$$\text{Gain (Total, Minat Mahasiswa)} = (0,89212128) - \left(\left(\frac{94}{110} x 0,70457671\right) + \left(\frac{16}{110} x 0\right)\right)$$

$$= 0,290028453$$

**Menghitung entropy dan gain Motivasi :**

$$\text{Entropy [Tinggi]} = \left(-\frac{49}{49} x \log_2 \left(\frac{49}{49}\right)\right) + \left(-\frac{0}{49} x \log_2 \left(\frac{0}{49}\right)\right)$$

$$= 0$$

$$\text{Entropy [Rendah]} = \left(-\frac{27}{61} x \log_2 \left(\frac{27}{61}\right)\right) + \left(-\frac{34}{61} x \log_2 \left(\frac{34}{61}\right)\right)$$

$$= 0,99047997$$

$$\text{Gain (Total, Motivasi)} = (0,89212128) - \left(\left(\frac{49}{110} x 0\right) + \left(\frac{61}{110} x 0,99047997\right)\right)$$

$$= 0,342855113$$

**Menghitung entropy dan gain Komunikasi :**

$$\text{Entropy [Terbuka]} = \left(-\frac{54}{71} x \log_2 \left(\frac{54}{71}\right)\right) + \left(-\frac{17}{71} x \log_2 \left(\frac{17}{71}\right)\right)$$

$$= 0,79410214$$

$$\text{Entropy [Tertutup]} = \left(-\frac{22}{39} x \log_2 \left(\frac{22}{39}\right)\right) + \left(-\frac{17}{39} x \log_2 \left(\frac{17}{39}\right)\right)$$

$$= 0,98811084$$

$$\text{Gain (Total, Komunikasi)} = (0,89212128) - \left(\left(\frac{71}{110} x 0,79410214\right) + \left(\frac{39}{110} x 0,98811084\right)\right)$$

$$= 0,029234238$$

**Menghitung entropy dan gain Media Pembelajaran :**

$$\text{Entropy [Cetak]} = \left(-\frac{35}{51} x \log_2 \left(\frac{35}{51}\right)\right) + \left(-\frac{16}{51} x \log_2 \left(\frac{16}{51}\right)\right)$$

$$= 0,89742719$$

$$\text{Entropy [Non Cetak]} = \left(-\frac{42}{59} x \log_2 \left(\frac{42}{59}\right)\right) + \left(-\frac{17}{59} x \log_2 \left(\frac{17}{59}\right)\right)$$

$$= 0,86630068$$

$$\text{Gain (Total, Media Pembelajaran)} = (0,89212128) - \left(\left(\frac{51}{110} x 0,89742719\right) + \left(\frac{59}{110} x 0,86630068\right)\right)$$

$$= 0,01138922$$

**Menghitung entropy dan gain Sarana Prasarana :**

$$\text{Entropy [Lengkap]} = \left(-\frac{54}{77} x \log_2 \left(\frac{54}{77}\right)\right) + \left(-\frac{23}{77} x \log_2 \left(\frac{23}{77}\right)\right)$$

$$= 0,87969758$$

$$\begin{aligned} Entropy \text{ [Tidak Lengkap]} &= \left( -\frac{22}{33} \times \log_2 \left( \frac{22}{33} \right) \right) + \left( -\frac{11}{33} \times \log_2 \left( \frac{11}{33} \right) \right) \\ &= 0,91829583 \end{aligned}$$

$$\begin{aligned} Gain \text{ (Total, Sarana Prasarana)} &= (0,89212128) - \left( \left( \frac{77}{110} \times 0,87969758 \right) + \right. \\ &\quad \left. \left( \frac{33}{110} \times 0,91829583 \right) \right) \\ &= 0,000844226 \end{aligned}$$

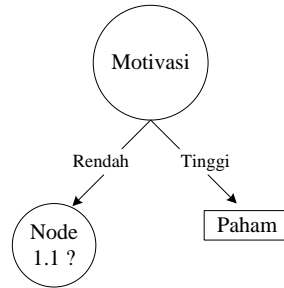
Berikut ini hasil perhitungan nilai *entropy* dan *gain* yang diuraikan pada Tabel 4.1.

**Tabel 4. Hasil Perhitungan Node 1**

NODE 1		Jumlah (S)	Paham	Tidak Paham	Entropy	Gain
total		110	76	34	0,89212128	
Minat Mahasiswa						0,290028453
	Ada	94	76	18	0,70457671	
	Tidak Ada	16	0	16	0	
Motivasi						<b>0,342855113</b>
	Tinggi	49	49	0	0	
	Rendah	61	27	34	0,99047997	
Komunikasi						0,029234238
	Terbuka	71	54	17	0,79410214	
	Tertutup	39	22	17	0,98811084	
Media Pembelajaran						0,01138922
	Cetak	51	35	16	0,89742719	
	Non Cetak	59	42	17	0,86630068	
Sarana Prasarana						0,000844226
	Lengkap	77	54	23	0,87969758	
	Tidak Lengkap	33	22	11	0,91829583	

Dari hasil perhitungan diatas, diperoleh nilai atribut tertinggi adalah Motivasi yang memiliki *gain* tertinggi **0,342855113**. Maka atribut Motivasi dipilih sebagai *node* akar. Nilai kelas atribut Tinggi adalah kosong, sehingga tidak perlu dilakukan perhitungan pada kelas atribut tersebut. Untuk kelas atribut Rendah belum diperoleh hasil antara keputusan paham dan tidak paham, maka perlu dilakukan perhitungan lebih lanjut. Maka pohon keputusan dari tabel diatas sebagai berikut :





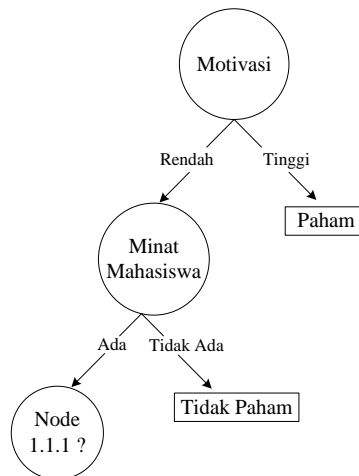
**Gambar 2. Pohon Keputusan Hasil Perhitungan Node 1**

Berdasarkan Gambar 4.1 diatas, atribut Motivasi digunakan sebagai *node* akar. Kelas atribut rendah belum diperoleh satu keputusan sehingga menghasilkan *node* cabang yaitu *Node 1.1*. Selanjutnya dilakukan perhitungan pada *Node 1.1* yaitu Motivasi = Rendah. Berikut ini hasil perhitungan dari kelas atribut Motivasi = Rendah yang ditunjukkan pada Tabel 5 berikut ini:

**Tabel 5. Hasil Perhitungan Node 1.1**

NODE 1.1		Jumlah (S)	Paham	Tidak Paham	Entropy	Gain
total		61	27	34	0,990479974	
Minat Mahasiswa						<b>0,252863116</b>
	Ada	45	25	18	0,99988063	
	Tidak Ada	16	0	16	0	
Komunikasi						0,030982508
	Terbuka	36	19	17	0,997772472	
	Tertutup	25	8	17	0,904381458	
Media Pembelajaran						0,00288487
	Cetak	27	11	16	0,975119065	
	Non Cetak	34	16	18	0,997502546	
Sarana Prasarana						0,00723294
	Lengkap	41	18	23	0,989245297	
	Tidak Lengkap	20	8	12	0,970950594	

Dari hasil perhitungan pada Tabel 5 tersebut, atribut yang menjadi *node* cabang dari Motivasi = Rendah adalah Minat Mahasiswa dengan nilai *gain* tertinggi sebesar 0,252863116. Nilai kelas atribut Tidak Ada menghasilkan keputusan Tidak Paham. Sedangkan nilai kelas atribut Ada belum diperoleh hasilnya, maka akan dilakukan perhitungan lanjut. Berikut ini digambarkan pohon keputusan berdasarkan tabel 5.



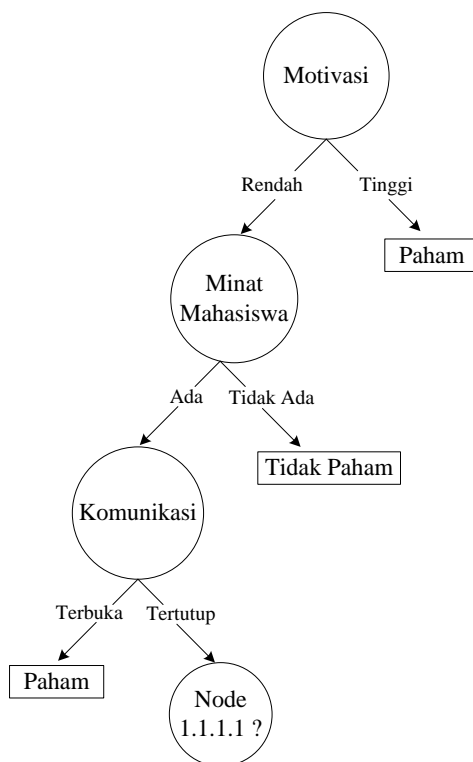
**Gambar 3. Pohon Keputusan Hasil Perhitungan Node 1.1**

Berikut ini hasil perhitungan Motivasi = Rendah dan Minat Mahasiswa = Ada yang ditunjukkan pada Tabel 6 berikut ini:

**Tabel 6. Hasil Perhitungan Node 1.1.1**

<b>NODE 1.1.1</b>		<b>Jumlah (S)</b>	<b>Paham</b>	<b>Tidak Paham</b>	<b>Entropy</b>	<b>Gain</b>
total		45	27	18	0,970950594	
Komunikasi						<b>0,00904948</b>
	Terbuka	28	18	10	0,940285959	
	Tertutup	17	9	8	0,997502546	
Media Pembelajaran						0,006004621
	Cetak	20	11	9	0,992774454	
	Non Cetak	25	16	9	0,942683189	
Sarana Prasarana						0,002512313
	Lengkap	31	18	13	0,981152234	
	Tidak Lengkap	14	9	5	0,940285959	

Dari hasil perhitungan pada Tabel 6 yang ditetapkan sebagai *node* cabang dari atribut Motivasi = Rendah dan Minat Mahasiswa = Ada adalah Komunikasi dengan *gain* tertinggi yaitu sebesar 0,00904948 dengan kelas atribut Terbuka diperoleh hasil Paham dan kelas atribut Tertutup belum diperoleh hasil keputusan. Oleh karena itu kelas atribut Tertutup dilakukan perhitungan selanjutnya. Berikut ini digambarkan pohon keputusan berdasarkan Tabel 6:



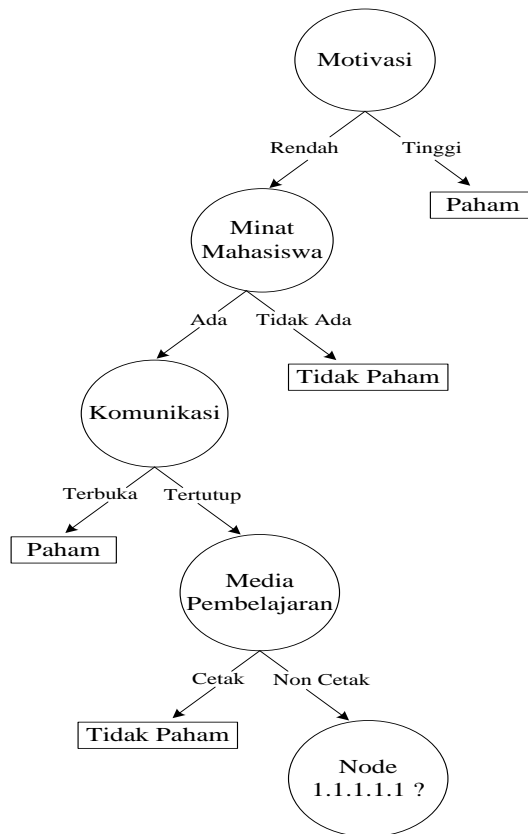
**Gambar 4. Pohon Keputusan Hasil Perhitungan Node 1.1.1**

Berikut ini hasil perhitungan dari kelas atribut Motivasi = Rendah dan Minat Mahasiswa = Ada dan Komunikasi = Tertutup yang ditunjukkan pada Tabel 7 berikut:

**Tabel 7. Hasil Perhitungan Node 1.1.1.1**

NODE 1.1.1.1		Jumlah (S)	Paham	Tidak Paham	Entropy	Gain
total		16	8	8	1	
Media Pembelajaran						<b>0,105843345</b>
	Cetak	7	2	5	0,863120569	
	Non Cetak	9	6	3	0,918295834	
Sarana Prasarana						0
	Lengkap	12	6	6	1	
	Tidak Lengkap	4	2	2	1	

Berdasarkan hasil perhitungan pada Tabel 7 tersebut, atribut yang menjadi *node* cabang dari Motivasi = Rendah dan Minat Mahasiswa = Ada dan Komunikasi Tertutup adalah Media Pembelajaran dengan nilai *gain* tertinggi sebesar 0,105843345 dengan kelas atribut Cetak diperoleh hasil Tidak Paham dan kelas atribut Non Cetak belum diperoleh hasil keputusan. Oleh karena itu kelas atribut Non Cetak dilakukan perhitungan selanjutnya. Berikut ini digambarkan pohon keputusan berdasarkan tabel 7:



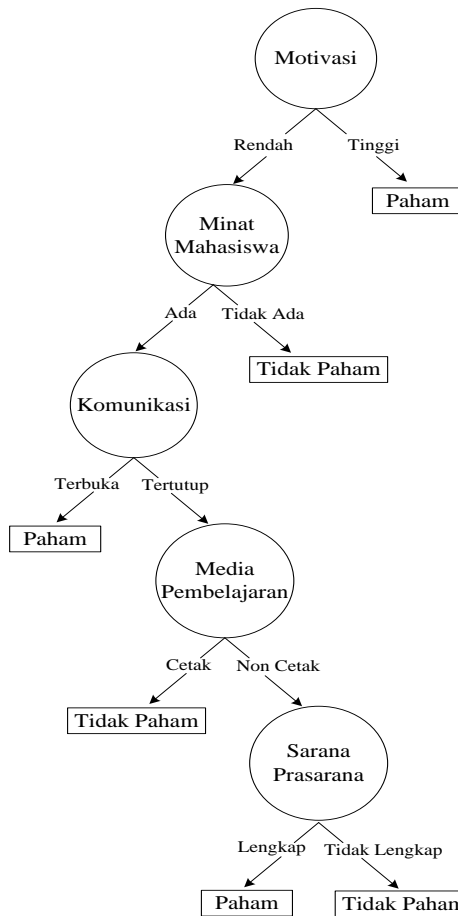
**Gambar 5. Pohon Keputusan Hasil Perhitungan Node 1.1.1.1**

Berikut ini hasil perhitungan selanjutnya dari Motivasi = Rendah dan Minat Mahasiswa = Ada dan Komunikasi = Tertutup dan Media Pembelajaran = Non Cetak yang ditunjukkan pada Tabel 8 berikut ini:

**Tabel 8. Hasil Perhitungan Node 1.1.1.1.1**

NODE 1.1.1.1.1		Jumlah (S)	Paham	Tidak Paham	Entropy	Gain
total		8	5	3	0,954434003	
Sarana Prasarana						<b>0,199203505</b>
	Lengkap	7	5	2	0,863120569	
	Tidak Lengkap	1	0	1	0	

Berdasarkan perhitungan Tabel 8 tersebut, telah memperoleh keputusan Paham pada kelas atribut Lengkap dan keputusan Tidak Paham pada kelas atribut Tidak Lengkap. Dengan demikian node 1.1.1.1.1 adalah node cabang terakhir yang terbentuk. Maka proses perhitungan telah selesai dilakukan kemudian pohon keputusan selanjutnya dapat dilihat pada gambar 6 berikut ini.



**Gambar 6. Pohon Keputusan Hasil Perhitungan Node 1.1.1.1.1**

Dari gambar 6 hasil perhitungan *Node* 1.1.1.1.1 diperoleh pohon keputusan atribut dengan hasil akhir Sarana Prasarana, Lengkap = Paham dan Tidak Lengkap = Tidak Paham.

## 2. Hasil Percobaan dan Pengujian

Pada tahap melalui pengujian menggunakan *software RapidMiner 5.3*. Hasil pengolahan data dengan model pohon keputusan sesuai dengan *software RapidMiner*.

## 3. Validasi Decision Tree

accuracy: 84.38%			
	true Paham	true Tidak Paham	class precision
pred. Paham	22	4	84.62%
pred. Tidak Paham	1	5	83.33%
class recall	95.65%	55.56%	

**Gambar 4.18. Nilai Akurasi Algoritma C4.5**

PerformanceVector		
PerformanceVector:		
accuracy: 84.38%		
ConfusionMatrix:		
True:	Paham	Tidak Paham
Paham:	22	4
Tidak Paham:	1	5

**Gambar 4.19. Nilai Performance Vector Algoritma C4.5**

Berdasarkan pengolahan data menggunakan *software rapidminer* didapat nilai akurasi sistem sebesar 84.38%. Dimana model yang telah dibentuk diuji tingkat akurasi nya, dengan memasukkan atau uji yang berasal dari data *training* dengan menggunakan *split validation* pada aplikasi *rapidminer 5.3* untuk menguji tingkat akurasi.

## KESIMPULAN DAN SARAN

Berdasarkan pembahasan sebelumnya dapat disimpulkan bahwa permasalahan dalam menentukan tingkat pemahaman mahasiswa pada matakuliah bahasa pemrograman dapat diselesaikan menggunakan teknik *data mining*, yaitu dengan Algoritma C4.5. Menghasilkan 6 *rules* dan Tingkat akurasi yang dihasilkan oleh metode tersebut adalah 84.38%. Dari perhitungan menggunakan Algoritma C4.5 maka didapatkan faktor yang paling dominan adalah Motivasi dengan nilai *gain* sebesar **0,342855113**.

## DAFTAR REFERENSI

- Anton, S. (2017). Implementasi algoritma Brute Force pada perancangan aplikasi Kamus Bahasa Sunda berbasis Android. *Jurnal Informatika Dan Bisnis*, 5(1), 1–12.
- Cynthia, E. P., & Ismanto, E. (2018). *METODE DECISION TREE ALGORITMA C.45 DALAM MENGLASIFIKASI DATA PENJUALAN BISNIS GERAJ MAKANAN CEPAT SAJI*. 3(3), 1–13.
- Junia, A., & Riandari, F. (2019). *Data mining untuk mengukur tingkat kepuasan peserta BPJS Ketenagakerjaan menggunakan Algoritma C4.5*. 1(1), 47–51.
- Mulyana, F. (2019). Universitas Sumatera Utara Skripsi.
- Novika, T., Poningsih, P., Okprana, H., Windarto, A. P., & Siahaan, H. (2021). Penerapan data mining klasifikasi tingkat pemahaman Siswa pada pelajaran Matematika. *Jurnal Media Informatika Budidarma*, 5(1), 9. <https://doi.org/10.30865/mib.v5i1.2498>
- Raharjo, M. R., & Windarto, A. P. (2021). Penerapan machine learning dengan konsep data mining rough set ( prediksi tingkat Pemahaman Mahasiswa terhadap Matakuliah ). *MEDIA INFORMATIKA BUDIDARMA*, 5, 317–326. <https://doi.org/10.30865/mib.v5i1.2745>
- Riandari, F., & Simangunsong, A. (2019). Penerapan algoritma C4.5 untuk mengukur tingkat kepuasan mahasiswa. *Jurnal Mantik Penusa*, 3(2), 1–7.
- Saputra, K. A., Hardinata, J. T., Lubis, M. R., Andani, S. R., & Saragih, I. S. (2020). Klasifikasi Algoritma C4.5 dalam penerapan tingkat kepuasan Siswa terhadap media

- pembelajaran online. *Media Online*, 1(3), 113–118.
- Setio, P. B. N., Saputro, D. R. S., & Bowo Winarno. (2020). Klasifikasi dengan pohon keputusan berbasis Algoritme C4.5. *PRISMA, Prosiding Seminar Nasional Matematika*, 3, 64–71.
- Sitorus, D. R., Windarto, A. P., Hartama, D., & Damanik, I. S. (2019). Penerapan klasifikasi C4.5 dalam meningkatkan Sistem pembelajaran Mahasiswa. *KOMIK (Konferensi Nasional Teknologi Informasi Dan Komputer)*, 3(1), 593–597. <https://doi.org/10.30865/komik.v3i1.1665>
- Sulastri, H., & Gufroni, A. I. (2017). Penerapan data mining dalam pengelompokan penderita Thalassaemia. *Jurnal Nasional Teknologi Dan Sistem Informasi*, 3(2), 299–305. <https://doi.org/10.25077/teknosi.v3i2.2017.299-305>
- Takalapeta, S. (2018). Penerapan data mining untuk menganalisis kepuasan konsumen menggunakan metode Algoritma C4.5. *J I M P - Jurnal Informatika Merdeka Pasuruan*, 3(3), 34–38. <https://doi.org/10.37438/jimp.v3i3.186>
- Zer, P. P. P. A. N. . F. I. R. ., Hartama, D., & Andani, S. R. (2019). Analisa faktor dominan Mahasiswa kesulitan memahami Bahasa Pemrograman menggunakan Metode C4.5. *Prosiding Seminar Nasional Riset Information Science (SENARIS)*, 1(September), 492. <https://doi.org/10.30645/senaris.v1i0.55>