

## Klasifikasi Tingkat Pemahaman Mahasiswa mengenai Teori Kepribadian Gustav Jung Menggunakan Algoritma C4.5

Dhea Indahsari<sup>1</sup>, Iqbal Maulana<sup>2</sup>, Aji Primajaya<sup>3</sup>

Universitas Singaperbangsa Karawang

<sup>1</sup>dhea.indahsari17084@student.unsika.ac.id, <sup>2</sup>iqbal.maulana@staff.unsika.ac.id,  
<sup>3</sup>aji.primajaya@staff.unsika.ac.id

(Naskah masuk: 03 Juni 2021, diterima untuk diterbitkan: 14 Juli 2021, Terbit: 28 Februari 2022)

### ABSTRAK

Pemahaman mahasiswa mengenai tipe kepribadian dapat digunakan sebagai tolak ukur dalam memahami dan mengeksplorasi diri. Pohon keputusan merupakan sebuah metode umum yang digunakan untuk memahami dalam pengambilan keputusan dan menginterpretasikan solusi dari permasalahan. Penelitian ini memiliki tujuan untuk mengklasifikasikan tingkat pemahaman mahasiswa mengenai kepribadian *introvert – ekstrovert* berdasarkan teori Gustav Jung dengan mengimplementasikan algoritma C4.5 dan mengevaluasi kinerja algoritma C4.5 dalam mengklasifikasikan tingkat pemahaman mahasiswa mengenai kepribadian dengan menggunakan *tools* Rapidminer. Algoritma C4.5 membentuk *decision tree* yang mudah dimengerti serta memberikan taraf akurasi yang dapat diterima dan juga efektif. Evaluasi terhadap model yang digunakan dilihat dari nilai *accuracy* dan nilai *kappa*. Berdasarkan hasil pengolahan data dengan mengimplementasikan algoritma C4.5 untuk mengklasifikasikan tingkat pemahaman mahasiswa terhadap kepribadian *introvert-ekstrovert* diperoleh bahwa mahasiswa memiliki tingkat pemahaman tinggi terhadap tipe kepribadian. Algoritma C4.5 dalam penelitian ini memiliki kinerja yang rendah dengan nilai *kappa* yang hanya mencapai 0,081. Akan tetapi, untuk tingkat akurasi mencapai nilai yang cukup yaitu sebesar 71,79%.

**Kata kunci:** *Algoritma C4.5, Data Mining, Decision Tree, Kepribadian*

### ABSTRACT

*Students' understanding of personality types serves as a benchmark for self-awareness and self-discovery. One of the commonly used methods for understanding decisions and interpreting solutions to problems is the decision tree. The purpose of this study is to classify the level of students' understanding about the personality of the introvert-extrovert based on Gustav Jung's theory using the C4.5 algorithm and evaluate the performance of the C4.5 algorithm in classifying the level of students' understanding of personality using tools Rapidminer. The C4.5 algorithm produces a decision tree that is easy to interpret and produces an acceptable and efficient level of accuracy. Evaluation of the model used seen from the value of accuracy and kappa value. Based on the results of data processing to implement the algorithm C4.5 to classify the level of students' understanding of the personality of the introvert-extrovert obtained that the students have a level of understanding high on personality types. The C4.5 algorithm in this study has a low performance with a kappa value of only 0.081. However, for the level of accuracy is sufficient, that value is 71.79% .*

**Keywords:** *C4.5 Algorithm, Data Mining, Decision Tree, Personality*

## 1. PENDAHULUAN

Kepribadian adalah suatu kesatuan yang terus bergerak dalam individu, terdiri dari banyak sistem fisik dan mental yang berbeda, serta merupakan faktor penentu dalam beradaptasi dengan lingkungan (Hamali, 2018). Carl Gustav Jung mengemukakan dalam penelitiannya bahwa manusia memiliki dua pandangan dalam menyalurkan perhatian dan energinya, yaitu *introvert* dan *ekstrovert* dimana keduanya bersifat kontradiktif (Susanto & Mudaim, 2017). Sikap *ekstrovert* mengalirkan energi kesadaran ke luar, menuju dunia objektif, sedangkan sikap *introvert* energinya ditujukan pada kesadaran internal menuju dunia subjektif. *ekstrovert* mengacu pada tipe orang yang menyukai dunia luar seperti *hang out*, menikmati interaksi, terlibat dalam aktivitas, fokus pada dunia luar dan pada tindakannya. Sedangkan seseorang dengan tipe kepribadian *introvert*, ia lebih menyukai dunia dalamnya dan hal tersebut berbanding terbalik dengan seorang *ekstrovert*. Mereka suka sendirian, membaca, dan menulis, dan mereka bisa bekerja sendiri (Susanto & Mudaim, 2017). Singkatnya, *introvert* sangat menikmati kedamaian dan ketenangan, sedangkan *ekstrovert* suka hidup di lingkungan yang melibatkan banyak interaksi atau dialog.

Menurut prayitno dalam penelitian Wijaya (2016) mahasiswa menghadapi realita proses perubahan jenjang pendidikan yang di dalamnya termasuk pematapan status dan identitas diri yang mendorong mahasiswa untuk meningkatkan kemampuan perannya. Kepribadian memberikan pengaruh terhadap kemampuan seseorang untuk beradaptasi dan menghadapi masalah (Pamungkas A., 2020). Pemahaman mahasiswa mengenai tipe kepribadian

*introvert* dan *ekstrovert* dapat digunakan sebagai tolak ukur dalam memahami dan mengeksplorasi diri berdasarkan kepribadiannya sendiri (Setiawati, F. A., dkk, 2015). Menurut Minggu dalam Listiawati (2015) Pemahaman adalah keterkaitan antara *planning* yang ada dengan informasi yang diperoleh. Konsep mengenai suatu konsep dapat dipahami jika mahasiswa dapat memahami makna, situasi, dan peristiwa yang diketahui, serta dapat mengasosiasikan konsep yang baru diperoleh dengan konsep yang telah diperoleh sebelumnya. Tujuh indikator yang dapat dikembangkan pada tingkat pemahaman tentang proses kognitif menurut Bloom dalam Minarto (2020) yaitu: 1) *interpreting* (interpretasi), keterampilan yang memungkinkan seseorang memperoleh ilmu / informasi mengenai objek tertentu dan dapat menafsirkannya kembali dengan cara lain. Misalnya interpretasi kata demi kata (paraphrase), kata demi ilustrasi, angka akan kata, kata ke angka, catatan dalam nada, dll; 2) *exemplifying* (mencontohkan), keahlian yang dimiliki seseorang dalam menyalurkan contoh mengenai sebuah gambaran yang telah dipelajari dalam tahapan edukasi. Contoh yang diberikan meliputi identifikasi definisi, karakteristik objek umum atau prinsip; 3) *classifying* (mengklasifikasikan), keahlian yang dimiliki seseorang untuk mengelompokkan suatu yang diawali dengan aktivitas seseorang yang diketahui pada suatu objek, hingga seseorang dapat menggambarkan kembali karakteristik dari suatu konsep, serta mengelompokkannya berdasarkan karakteristik yang telah ditemui oleh seseorang tersebut; 4) *summarizing* (meringkas), kemampuan yang dimiliki oleh seseorang dalam mengembangkan penyampaian yang mampu

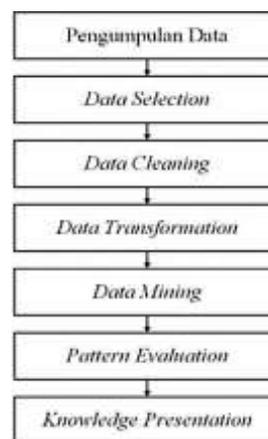
menggambarkan isi data/ topik secara totalitas dalam wujud ringkasan/ resume ataupun abstrak; 5) *inferring* (menyimpulkan), kemampuan yang dimiliki seseorang dalam menciptakan pola cerminan topik yang diberikan. Kegiatan ini adalah aktivitas lanjutan dari aktivitas pembuatan ringkasan atau abstrak dengan karakteristik yang relevan dan berkaitan dengan jelas antara keduanya; 6) *comparing* (membandingkan), keterampilan yang terdapat dalam diri seseorang untuk mengetahui persamaan serta perbandingan antara 2 objek ataupun lebih, peristiwa, inspirasi, permasalahan, ataupun suasana, selaku penentuan bagaimana peristiwa bisa terjalin dengan baik; 7) *explaining* (menjelaskan), keahlian yang terdapat dalam diri seseorang sehingga seseorang tersebut dapat meningkatkan serta memakai pemicu ataupun pengaruh objek yang dialokasikan.

Penelitian ini memiliki tujuan untuk mengklasifikasikan tingkat pemahaman mahasiswa mengenai tipe kepribadian berdasarkan teori Carl Gustav Jung dengan mengimplementasikan algoritma C4.5. Algoritma C4.5 membentuk *decision tree* yang tidak sulit untuk dimengerti serta memberikan taraf akurasi yang dapat diterima dan efektif (Asmirayani, 2019). Dalam hal ini, penulis mencoba menggunakan nilai *kappa* untuk mendukung taraf akurasi pada hasil evaluasi dari kinerja algoritma C4.5 dalam mengklasifikasikan tingkat pemahaman mahasiswa mengenai teori kepribadian Gustav Jung.

## 2. METODELOGI

Penelitian ini dilaksanakan dengan menggunakan metodologi KDD (*Knowledge Discovering in Databases*) untuk menentukan tingkat pemahaman mahasiswa mengenai tipe kepribadian *introvert* dan *ekstrovert*. Data yang akan

diolah diperoleh dari metode kuesioner terhadap beberapa sampel untuk memudahkan peneliti dalam menghitung hasil uji. Data - data yang berpengaruh dalam penelitian diolah lebih lanjut dengan algoritma C4.5 sehingga data tersebut menghasilkan informasi mengenai tingkat pemahaman mahasiswa menjadi tiga tingkat yaitu kurang, cukup dan baik. Tahapan dari penelitian dapat dilihat pada gambar 2.1.



Gambar 2.1 Tahapan Penelitian

### 2.1 Data Penelitian

Penelitian ini menggunakan data primer, dimana data tersebut dikumpulkan secara eksklusif dari hasil kuesioner terhadap subjek penelitian. Dalam pelaksanaannya, kuesioner ini ditujukan kepada mahasiswa Universitas Singaperbangsa yang memenuhi kriteria yang ditetapkan.

Jumlah mahasiswa aktif Universitas Singaperbangsa Karawang berdasarkan Pangkalan Data Perguruan Tinggi (PDDikti) adalah sebanyak 18.766 mahasiswa. Teknik dalam pengambilan sampel untuk penelitian ini ditentukan dengan metode Slovin. Berdasarkan perhitungan yang telah dilakukan dengan metode tersebut, maka jumlah sampel yang terlibat dalam penelitian ini adalah 392 siswa.

Penelitian memerlukan variabel penelitian agar peneliti dapat mengetahui

apa yang sedang dipelajarinya. Variabel yang diimplementasikan pada penelitian ini yaitu tipe kepribadian dan tingkat pemahaman, dimana menurut Altemeyer dalam Zaduqisty (2019) menyatakan bahwa Kepribadian merupakan salah satu faktor yang mempengaruhi pemahaman. Oleh karena itu, variabel yang dimanfaatkan pada penelitian ini adalah sebagai berikut :

Tabel 2.1 Variabel Penelitian

Variabel	Parameter	Hasil Ukur
Tipe Kepribadian	Kepribadian menurut C. G. Jung	<i>Introvert</i>
		<i>Ekstrovert</i>
Tingkat Pemahaman	Tingkat pemahaman menurut Bloom : <i>Interpreting</i> <i>Exemplifying</i> <i>Classifying</i> <i>Summarizing</i> <i>Inferring</i> <i>Comparing</i> <i>Explaining</i>	Baik jika nilainya $\geq 76-100\%$
		Cukup jika nilainya $60-75\%$ .
		Kurang jika nilainya $\leq 60\%$
		(Arikunto, 2013)

Kuesioner tingkat pemahaman mahasiswa melingkupi 7 parameter diantaranya *interpreting*, *exemplifying*, *classifying*, *summarizing*, *inferring*, *comparing* dan *explaining* yang masing-masing berisi 2 soal. Uji validitas dan reliabilitas perlu dilakukan terhadap kuesioner yang akan dijadikan instrumen dalam penelitian. Uji validitas dan reliabilitas merupakan aspek krusial yang harus diperhatikan oleh peneliti bahwa instrumen penelitian tersebut bekerja dengan baik dan hasilnya dapat digunakan untuk menghasilkan data yang baik (Arifin, 2017). Berdasarkan hasil dari uji validitas dengan menggunakan aplikasi SPSS, diperoleh bahwa terdapat satu pernyataan yang tidak valid dalam kuesioner yang diimplementasikan sebagai instrumen dalam penelitian ini sehingga pernyataan

pada kuesioner yang tidak valid tersebut akan gugur sehingga tidak akan diimplementasikan sebagai pernyataan dalam instrumen penelitian karena sudah terwakili dengan butir pernyataan kuesioner yang lain.

Reliabilitas menunjukkan bahwa instrumen ini sudah sangat baik sehingga dapat diandalkan sebagai alat pengumpulan data (Arikunto, 2013). Pengujian reliabilitas penelitian ini dilakukan dengan menggunakan software SPSS.

Cronbach's Alpha	N of Items
,677	13

Gambar 2.2 Hasil uji reabilitas

Hasil uji reliabilitas terhadap variabel dengan *Cronbach Alpha* seperti yang terlihat dalam gambar 2.2 di atas menunjukkan bahwa nilai *Cronbach's Alpha*  $> 0,6$ . Oleh karena itu, diperoleh bahwa instrumen yang diimplementasikan dalam penelitian tersebut konsisten dalam mengumpulkan data.

## 2.2 Knowledge Discovery in Databases

Han dan Kamber (2012) menggunakan istilah *Knowledge Discovery in Database* (KDD) dalam menjelaskan proses penggalian informasi tersembunyi dari kumpulan *big data*. Namun, meskipun konsep dari kedua istilah ini berbeda, namun kedua istilah tersebut saling terkait dan salah satu langkah dalam proses KDD adalah *data mining*. Langkah-langkah penting dalam proses KDD menurut Siregar dan Puspabhuana (2017), meliputi tahapan berikut: 1) *Data Selection*, adalah tahap dalam memilih dan serta penganalisisan data operasi. Data yang telah dipilih akan tersimpan di basis data yang berbeda-beda; 2) *Data Cleaning*, tahapan dalam menghilangkan data imitasi, memeriksa ketidakkonsistenan data dan memperbaiki data yang salah (seperti

kesalahan penulisan). Dalam banyak kasus, isi data tidak lengkap, seperti data hilang, data invalid, atau kesalahan dalam pengetikan. Selain itu, menghilangkan beberapa atribut data yang tidak ada hubungannya dengan asumsi data mining; 3) *Data Integration*, proses penambahan data yang sudah ada ke data atau data terkait, atau proses yang disebut informasi, adalah penggabungan data dari basis data yang berbeda ke dalam basis data baru yang dibutuhkan KDD; 4) *Data Transformation*, proses konversi data menjadi data yang sesuai untuk tahapan selanjutnya, yaitu *data mining* dengan format tertentu. Misalnya, beberapa metode standar (seperti metode asosiasi dan klasifikasi) hanya dapat menerima entri data kategoris; 5) *Data Mining*, proses menemukan pengetahuan berupa informasi atau pola yang berbeda dengan mengimplementasikan teknik, metode, atau algoritma tertentu. 6) *Pattern Evaluation*, identifikasi model yang sangat unik dalam penggalian data. Pada tahap ini, hasil dari teknik *data mining* dievaluasi dalam bentuk model unik dan model prediktif untuk menilai dugaan yang ada memang terpenuhi; 7) *Knowledge Presentation*, representasi pengetahuan menggambarkan sampel informasi yang diambil dari tahapan penggalian data. Visualisasi ini membantu mengomunikasikan hasil penambangan data dengan cara yang mudah dipahami.

### 2.3 Algoritma C4.5

Menurut Elisa (2017) algoritma C4.5 merupakan algoritma yang terkenal dalam mengklasifikasikan data menggunakan atribut numerik dan kategorik. Proses klasifikasi menghasilkan aturan yang dapat digunakan dalam meramalkan nilai dari atribut dengan tipe diskrit *record* baru. Algoritma C4.5 sendiri merupakan evolusi dari algoritma ID3 yang pengembangannya

bertujuan untuk mengatasi data yang hilang dan dapat menangani data yang terus menerus dan terpangkas.

Algoritma C4.5 didasarkan pada algoritma yang merepresentasikan pohon keputusan. Secara generik tahapan algoritma C4.5 dalam membentuk sebuah *decision tree* adalah sebagai berikut (Elisa, 2017): 1) Seleksi atribut sebagai *root*, 2) Buat cabang untuk setiap nilai; 3) Bagi kasus pada cabang; 4) Proses terus berulang pada cabang hingga semua kasus dalam cabang memiliki kelas yang tidak memiliki perbedaan.

Untuk menentukan atribut yang menjadi akar dari pohon keputusan, berdasarkan pada nilai *gain* yang lebih tinggi dari atribut - atribut yang ada. Persamaan yang digunakan dalam menghitung *gain* rumus seperti yang tertera pada persamaan 1 berikut (Kusrini, 2019).

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

Keterangan :

S : Kompilasi kasus

A : Atribut

|S<sub>i</sub>| : Jumlah kasus pada partisi ke-i

|S| : Jumlah kasus dalam S

Sementara itu, nilai entropi dapat dihitung berdasarkan pada persamaan 2 berikut.

$$\text{Entropy (S)} = \sum_{i=1}^n - p_i * \log_2 p_i \quad (2)$$

Keterangan :

n : Jumlah partisi s

p<sub>i</sub> : proporsi dari S<sub>i</sub> terhadap S

*Entropy* merupakan ukuran berdasarkan teori informasi yang dapat menentukan sifat impurity serta keseragaman dari himpunan data. Setelah nilai entropy diperoleh, maka proses selanjutnya menghitung nilai dari *information gain* dari setiap atribut (Sukma, A.R., dkk, 2019). Dan *gain* merupakan data yang tidak sesuai dengan informasi dari data - data, baik Melalui partisipasi dalam informasi atau kumpulan data yang

diperoleh dari observasi (Harryanto & Hansun, 2017).

### 3. HASIL DAN PEMBAHASAN

#### 3.1 Pengumpulan Data

Pengumpulan data dilakukan dengan mendistribusikan kuesioner kepada responden dengan mengimplementasikan metode *simple random sampling*. Pengambilan sampel dari anggota populasi dikatakan sederhana karena dilakukan secara *random* dengan tidak memperhatikan struktur dari populasi yang ada (Sugiyono, 2011). Data yang diperoleh dari kuesioner yang telah didistribusikan dan dikumpulkan akan digabungkan menjadi satu informasi lengkap untuk dapat diproses ke tahap selanjutnya. Jumlah data yang didapatkan secara keseluruhan yaitu mencapai 400 data mahasiswa dengan karakteristik responden terdapat pada tabel berikut:

Tabel 3.1 Karakteristik Responden

Atribut	Nilai	Persentase
Semester	2	15,5%
	4	11,3%
	6	18,5%
	8	54,8%
Fakultas	Agama Islam	5%
	Ekonomi	13,8%
	Hukum	1,5%
	Ilmu Kesehatan	2,1%
	Ilmu Komputer	26,7%
	Ilmu Sosial dan Ilmu Politik	9%
	Keguruan dan Ilmu Pendidikan	29,5%
	Pertanian	4,5%
Teknik	7,9%	

Berdasarkan pada tabel 3.1 di atas, Karakteristik responden dalam penelitian ini diklasifikasikan berdasarkan fakultas dan tingkatan semester responden. Dan berdasarkan perhitungan terhadap sampel responden, diperoleh bahwa responden terbanyak adalah dari semester 8 dengan persentase mencapai 54,8%, disusul oleh semester 6 dengan presentase 18,5% ,

selanjutnya semester 2 dengan 15,5% dan semester 4 dengan 11,3%. Dari tabel tersebut dapat mengidentifikasi bahwa responden dengan partisipasi terbanyak adalah mahasiswa dari fakultas keguruan dan ilmu pendidikan dengan persentase mencapai 29,5%.

#### 3.2. Data Selection

Proses *data selection* dilakukan untuk menghapus nilai dan atribut yang tidak diperlukan dalam penelitian. Pada penelitian ini, ada beberapa nilai yang dihapus dikarenakan responden tidak termasuk kedalam kriteria responden yang dibutuhkan dalam penelitian, sehingga jumlah data menjadi 392 *record*. Selain itu, dari beberapa atribut yang diperoleh dari hasil kuesioner, hanya ada 9 atribut yang digunakan dalam klasifikasi. Adapun atribut yang digunakan pada penelitian ini meliputi *classifying, comparing, exempling, explaining, inferring, interpreting, summarizing*, tipe kepribadian dan *score*.

#### 3.3 Data Cleaning

Pada tahap ini, data dibersihkan dari *noise* ataupun data yang inkonsisten. Namun, karena data berasal dari kuesioner yang bersifat pernyataan tertutup, maka pada tahap ini tidak diberikan perlakuan apapun terhadap data karena tidak terdapat *missing value* ataupun data yang inkonsisten.

Name	Type	Missing
✓ Tipe Kepribadian	Binominal	0
✓ Interpreting	Binominal	0
✓ Exemplifying	Binominal	0
✓ Classifying	Binominal	0
✓ Summarizing	Binominal	0
✓ Inferring	Binominal	0
✓ Comparing	Binominal	0
✓ Explaining	Binominal	0
✓ Tingkat Pemahaman	Nominal	0

Gambar 3.1 Missing Value

Sebagaimana yang telah ditunjukkan pada gambar 3.1, bahwa *dataset* tingkat pemahaman mahasiswa yang akan diolah dengan bantuan *tools* Rapidminer tersebut tidak memiliki *missing value* sehingga data tersebut dapat langsung diolah untuk tahap selanjutnya.

### 3.4 Data Transformation

Transformasi data merupakan tahapan dalam mengkonversi data kedalam format yang sesuai untuk prosedur lebih lanjut. Berdasarkan kuesioner yang telah didapatkan, tingkat pemahaman mahasiswa mengenai tipe kepribadian dibagi menjadi 3 (tiga) penilaian seperti pada Tabel 3.2 berikut.

Tabel 3.2 Transformasi tingkat pemahaman

Range	Keterangan
76-100 %	Baik

Tabel 3.3 Perhitungan nilai *entropy* dan *gain*

	Jumlah Kasus	Kurang	Cukup	Baik	Entropy	Gain	
Total	353	23	86	244	1,121313		
Interpreting	Kurang	9	1	5	3	1,3516441	0,0103832
	Cukup	344	22	81	241	1,1046319	
Exemplifying						0,008795	

60–75 %.	Cukup
<60 %	Kurang

### 3.5 Data Mining

Tahap *data mining* dilakukan dengan memilih metode algoritma untuk menemukan pola atau informasi dalam mengklasifikasikan tingkat pemahaman mahasiswa mengenai tipe kepribadian *introvert* dan *ekstrovert* berdasarkan teori Carl Gustav Jung. Data tersebut terdiri atas 392 *record* dengan 9 atribut. Atribut tingkat pemahaman digunakan sebagai *class target*, sehingga atribut data menjadi 8 yaitu *interpreting*, *exemplifying*, *classifying*, *summarizing*, *inferring*, *comparing*, *explaining* dan tipe kepribadian. Algoritma C4.5 merupakan model yang akan diimplementasikan dalam penggalian data penelitian ini. Langkah pertama yaitu membagi data yang diperoleh menjadi data *training* dan data *testing* dengan perbandingan 9:1, sehingga jumlah data latih menjadi 353 *record* dengan data uji 39 *record*. Perbandingan tersebut dipilih dikarenakan menghasilkan nilai evaluasi yang lebih baik dibandingkan dengan rasio perbandingan lainnya pada penelitian ini. Setelah semua data yang diproses telah dibagi, proses klasifikasi dilakukan dengan membangun pohon keputusan sebagai keluaran.

Pohon keputusan pada algoritma C4.5 dikembangkan dengan sebelumnya dengan menghitung nilai *entropy* dan *gain* dari data. Hasil yang diperoleh dari perhitungan nilai *entropy* dan *gain* terhadap data latih dapat dilihat pada tabel 3.3 .

Classifying	Kurang	3	0	1	2	0	0,0096126
	Cukup	350	23	85	242	1,1220538	
Summarizing	Kurang	22	4	3	15	1,2158749	0,0245192
	Cukup	331	19	83	229	1,1047763	
Inferring	Kurang	11	4	4	3	1,5726236	0,0011672
	Cukup	342	19	82	241	1,0814893	
Comparing	Kurang	94	5	25	64	1,110907	0,0045408
	Cukup	259	18	61	180	1,1234986	
Explaining	Kurang	53	4	17	32	1,2470361	0,0039953
	Cukup	300	19	69	212	1,0937588	
Tipe Kepribadian	Kurang	145	12	38	95	1,2035154	0,0007612
	Cukup	208	11	48	149	1,0572279	
	<i>Introvert</i>	258	18	62	178	1,1317723	
	<i>Ekstrovert</i>	95	5	24	66	1,090079268	

Berdasarkan hasil yang diperoleh dari perhitungan *entropy* dan *gain* pada tabel di atas, nilai *gain* tertinggi terdapat pada atribut *Summarizing*. Oleh karena itu, atribut *Summarizing* akan menjadi *node root* dari pohon keputusan. Perhitungan terus berlanjut hingga memenuhi kondisi sebagai berikut : 1) Semua data partisi memiliki kelas yang sama; 2) Tidak ada lagi atribut untuk memisahkan data; 3) Tidak ada lagi data dalam partisi.

Suatu teknik klasifikasi dikatakan baik jika menghasilkan dugaan dari model tersebut dengan nilai akurasi yang tinggi. Selain berdasarkan tingkat akurasi, performa algoritma C4.5 dapat diukur berdasarkan nilai *kappa*. Hasil dari evaluasi kinerja algoritma C4.5 dalam mengklasifikasikan tingkat pemahaman tipe kepribadian *introvert* dan *ekstrovert* berdasarkan teori Gustav Jung termasuk ke dalam *poor classification* karena nilai *kappa* dari kinerja algoritma C4.5 yang digunakan dalam penelitian ini hanya mencapai 0,081. Akan tetapi, untuk tingkat akurasi yang diperoleh mencapai nilai yang termasuk dalam klasifikasi lumayan yaitu sebesar 71,79%. Hasil evaluasi kinerja algoritma C4.5 dengan

bantuan *tools* Rapidminer terdapat pada gambar 3.2 berikut.

```

PerformanceVector:
accuracy: 71.79%
ConfusionMatrix:
True:   Cukup   Baik   Kurang
Cukup:  1         1     0
Baik:   8         27    2
Kurang: 0         0     0
kappa:  0.081
ConfusionMatrix:
True:   Cukup   Baik   Kurang
Cukup:  1         1     0
Baik:   8         27    2
Kurang: 0         0     0
    
```

Gambar 3.2 Hasil Evaluasi

Dari gambar 3.2 diatas, berdasarkan pengolahan data dengan menggunakan bantuan *tools* Rapidminer diperoleh bahwa jumlah *record* pada data uji yang diklasifikasikan tepat oleh algoritma C4.5 sebanyak 28 *record* dari 39 *record*. Maka diperoleh bahwa persentase dari tingkat akurasi terhadap klasifikasi tingkat pemahaman mahasiswa dengan menggunakan algoritma C4.5 sebesar 71,79% dengan nilai *kappa* 0,081.

### 3.6 Knowledge Presentation

*Knowledge presentation* yang didapatkan dari perhitungan algoritma C4.5 dapat dipresentasikan dalam dua bentuk, yaitu dalam bentuk pohon keputusan dan *rule IF-THEN*. Pohon keputusan yang dibentuk oleh Rapidminer dapat dilihat pada gambar 3.3.



Gambar 3.3 *Decision tree*

Berdasarkan gambar 3.3, model yang dihasilkan merupakan pohon keputusan dengan 15 *rule* yang dimana atribut *summarizing* sebagai *root node* dari pohon keputusan tersebut. Selain itu, Rapidminer juga menyediakan *tool* yang dapat menampilkan hasil *decision tree* berupa *rule text IF-THEN* seperti yang ada pada gambar 3.4 berikut:

```

Tree
Summarizing = Cukup
|
|_ Interpreting = Cukup
|   |
|   |_ Explaining = Cukup
|   |   |
|   |   |_ Comparing = Cukup: Baik (Cukup=0, Baik=0, Kurang=1)
|   |   |_ Inferring = Cukup
|   |   |   |
|   |   |   |_ Explaining = Cukup: Baik (Cukup=0, Baik=0, Kurang=0)
|   |   |   |_ Explaining = Kurang
|   |   |   |   |
|   |   |   |   |_ Tipe Kepribadian = Ekstrovert: Cukup (Cukup=0, Baik=0, Kurang=0)
|   |   |   |   |_ Tipe Kepribadian = Introvert: Baik (Cukup=0, Baik=0, Kurang=1)
|   |   |   |   |_ Inferring = Kurang
|   |   |   |   |   |
|   |   |   |   |   |_ Tipe Kepribadian = Ekstrovert: Cukup (Cukup=0, Baik=0, Kurang=0)
|   |   |   |   |   |_ Tipe Kepribadian = Introvert: Baik (Cukup=0, Baik=0, Kurang=0)
|   |   |   |   |   |_ Explaining = Kurang
|   |   |   |   |   |   |
|   |   |   |   |   |   |_ Explaining = Kurang
|   |   |   |   |   |   |_ Comparing = Kurang
|   |   |   |   |   |   |   |
|   |   |   |   |   |   |   |_ Comparing = Kurang: Baik (Cukup=0, Baik=0, Kurang=0)
|   |   |   |   |   |   |   |_ Explaining = Kurang: Cukup (Cukup=0, Baik=0, Kurang=0)
|   |   |   |   |   |   |   |_ Interpreting = Kurang: Cukup (Cukup=0, Baik=0, Kurang=0)
|   |   |   |   |   |   |   |_ Summarizing = Kurang
|   |   |   |   |   |   |   |   |
|   |   |   |   |   |   |   |   |_ Classifying = Cukup
|   |   |   |   |   |   |   |   |   |
|   |   |   |   |   |   |   |   |   |_ Comparing = Cukup: Baik (Cukup=0, Baik=0, Kurang=0)
|   |   |   |   |   |   |   |   |   |_ Comparing = Kurang: Cukup (Cukup=0, Baik=0, Kurang=0)
|   |   |   |   |   |   |   |   |   |_ Explaining = Kurang: Cukup (Cukup=0, Baik=0, Kurang=0)
|   |   |   |   |   |   |   |   |   |_ Classifying = Kurang: Baik (Cukup=0, Baik=0, Kurang=0)

```

Gambar 3.4 *Rule text decision tree*

Gambar 3.4 di atas merupakan *rule text IF-THEN* dari pohon keputusan yang terbentuk dari hasil perhitungan menggunakan algoritma C4.5 dengan bantuan *tools* Rapidminer. *Rule IF-THEN* tersebut merupakan ekstraksi dari pohon keputusan dengan penelusuran dari akar sampai ke daun. *If* ditentukan oleh *node* dan cabang, sedangkan *then* merupakan nilai daun pada pohon keputusan.

#### 4. KESIMPULAN

Penelitian terhadap tingkat pemahaman mahasiswa mengenai tipe kepribadian *introvert* dan *ekstrovert* dengan menggunakan metodologi *knowledge discovery of databases (KDD)* yang dimana tahapannya meliputi *data cleaning*, *data selection*, *data transformation*, *data mining*, *pattern evaluation* dan *knowledge presentation*. Algoritma C4.5 diterapkan pada tahap data mining untuk menemukan pola tingkat pemahaman mahasiswa terhadap tipe kepribadian *introvert* dan *ekstrovert* berdasarkan teori Carl Gustav Jung. Hasil klasifikasi dengan menggunakan pohon keputusan yang telah melalui tahap evaluasi menyatakan bahwa tingkat pemahaman mahasiswa terhadap teori kepribadian Gustav Jung memiliki tingkat pemahaman tinggi. Kinerja algoritma C4.5 dalam mengklasifikasikan tingkat pemahaman mahasiswa mengenai teori kepribadian Gustav Jung termasuk ke dalam *poor classification* karena nilai *kappa* kinerja dari algoritma C4.5 dalam penelitian ini hanya mencapai 0,081. Akan tetapi, untuk tingkat akurasi mencapai nilai yang cukup sebesar 71,79%. Untuk pengembangan penelitian yang telah dilakukan, peneliti memberikan saran agar menggunakan lebih banyak data untuk meningkatkan akurasi pohon keputusan dan lakukan perbandingan algoritma C4.5

dengan model algoritma yang lainnya untuk memperoleh tingkat akurasi yang lebih akurat dan tinggi untuk mendukung pengujian data yang ada.

#### DAFTAR PUSTAKA

- Arifin, Z., 2017. Kriteria Instrumen dalam suatu Penelitian. *Jurnal THEOREMS (The Original Research of Mathematics)*, pp. 28-36.
- Arikunto, S., 2013. *Dasar-Dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara.
- Asmirayani, 2019. Penerapan Data Mining Dengan Algoritma C4.5 untuk Kalkulasi Data Dalam Al-Qur'an dan Terjemahan. *Kumpulan Jurnal Ilmu Komputer*, pp. 152-161.
- Elisa, E., 2017. Analisa dan Penerapan Algoritma C4.5 Dalam Data Mining Untuk Mengidentifikasi Faktor-Faktor Penyebab Kecelakaan Kerja Kontruksi PT.Arupadhatu Adisesanti. *JOIN*, pp. 36-41.
- Hamali, S., 2018. Kepribadian dalam Teori Sigmound Freud dan Nafsiologi dalam Islam. *Al-Adyan*, pp. 285-302.
- Han, J., Kamber, M. & Pei, J., 2012. *Data Mining : Concept and Techniques Third Edition*. United States of America: Morgan Kaufmann.
- Harryanto, F. F. & Hansun, S., 2017. Penerapan Algoritma C4.5 untuk Memprediksi Penerimaan Calon Pegawai Baru di PT WISE. *Jatisi*, pp. 95 - 103.
- Listiawaty, E., 2015. Pemahaman Mahasiswa Calon Guru Pada Konsep Grup. *Jurnal APOTEMA*, pp. 76 - 86.
- Minarto, 2020. Pemunculan Tingkat Kesulitan Soal Pada Tes Penjurusan Menggunakan Revised Bloom Taxonomi (RBT) Di SMAN 31 Bangorejo dengan Aplikasi Wingen3. *INCARE*, 01(01), pp. 17 - 27.
- Pamungkas, A., 2020. Kepribadian Ekstrovert- Introvert dan Kecemasan Mahasiswa pada masa Pandemi Covid-19. *Syams : Jurnal Studi Keislaman*, pp. 36 - 42.
- Setiawan, A., 2020. Algoritma Decision Tree C4.5 Untuk Klasifikasi Tingkat Pengetahuan Masyarakat Terhadap Pemanfaatan Bank Sampah Di Kabupaten Banjarmasin Selatan. *Technologia*, pp. 147-151.
- Siregar, A. M. & Puspabhuana, A., 2017. *DATA MINING: Pengolahan Data Menjadi Informasi dengan Rapid Miner*. s.l.:CV Kekata Group.
- Sugiyono, 2011. *Metode Penelitian Bisnis Pendekatan Kuantitatif Dan Kualitatif, Dan R&D*. Bandung: Alfabeta.
- Sukma, A. R., Halfis, R. & Hermawan, A., 2019. Klasifikasi Channel Youtube Indonesia menggunakan Algoritma C4.5. *Jurnal Teknik Komputer*, pp. 21 - 28.
- Susanto, E. & Mudaim, 2017. Pengembangan Inventori MBTI Sebagai Alternatif Instrumen Pengukuran Tipe Kepribadian. *Indonesia Journal Of Educational Conseling*, pp. 41 - 52.
- Wahyudi, R., 2020. Uji Validitas dan Reabilitas dengan Pendekatan Konsistensi Internal Kuesioner Pembukaan Program Studi Statistika FMIPA Universitas

Bengkulu. *Jurusan Matematika FMIPA Universitas Bengkulu*, pp. 1-10.

Wijaya, R. S., 2016. Perbandingan Penyesuaian Diri Mahasiswa Berkepribadian Ekstrovert dan Introvert. *Jurnal Penelitian Tindakan Bimbingan dan Konseling*, pp. 1 - 6.

Zaduqisti, E., 2019. Pemahaman Konsep Jihad Ditinjau Dari Identifikasi Kelompok, Persepsi Ancaman Antar-Kelompok, Dan Kepribadian. *Indonesian Psychological Research*, pp. 47-66.