

IMPLEMENTASI METODE DECISION TREE DAN ALGORITMA C4.5 UNTUK KLASIFIKASI KEPRIBADIAN MASYARAKAT

Krisna Ferdinan Leo Simanjuntak¹, Annita Carolina Br Barus², Anita³

^{1,2,3} Fakultas Teknologi dan Ilmu Komputer, Universitas Prima Indonesia

Sekip, Simpang Seikambing, Medan, Indonesia, 20111

Email : ¹ krisnasimanjuntak1708@gmail.com, ² annitacarolina0@gmail.com, ^{3*} anitayakub_pilchan@yahoo.com

Abstract

Personality is the overall attitude, feeling, temperament, characteristics and behavior of a person. so far in the village of Ujung Serdang Tanjung Morawa there is no personality classification in the local community, due to the lack of attention from the local government in analyzing the personality of the community. As a result of this problem, many people like to criticize the government's performance, do not obey the rules from the government and go against the rules made by the local government or even resort to violence, causing unrest for other communities. This study aims to make the implementation of data mining in helping the local government to analyze the personality of the community. The attributes used to analyze are gender, age, education and by asking 40 questions in the form of a questionnaire. The system was developed using the PHP version 7 programming language and applied the decision tree method and the C4.5 algorithm to analyze the personality of the community. Based on the results obtained from 40 community data that were tested, the information obtained was sanguin 17, choleric 6, melancholic 4, and phlegmatic 13.

Keywords: Decision Tree, C.45, Classification, Personality.

Abstrak

Kepribadian adalah keseluruhan sikap, perasaan, tempramen, ciri-ciri khas dan perilaku seseorang. selama ini di desa Ujung Serdang Tanjung Morawa belum adanya klasifikasi kepribadian pada masyarakat setempat, di sebabkan kurangnya perhatian pemerintah setempat dalam menganalisis kepribadian masyarakat tersebut. Akibat dari persoalan ini banyak masyarakat yang suka mengkritik kinerja pemerintahan, tidak taat akan aturan dari pemerintah dan melawan aturan yang di buat pemerintah setempat atau bahkan sampai melakukan kekerasan sehingga menimbulkan keresahan terhadap masyarakat lain. Penelitian ini bertujuan untuk membuat implementasi data mining dalam membantu pemerintah setempat untuk menganalisis kepribadian masyarakat tersebut. Atribut yang di gunakan untuk menganalisis ialah jenis kelamin, usia, pendidikan dan dengan mengajukan 40 pertanyaan dalam bentuk kuisioner. Sistem yang di kembangkan dengan menggunakan bahasa pemrograman PHP versi 7 dan menerapkan metode *decision tree* dan *algoritma C4.5* untuk menganalisis kepribadian masyarakat tersebut. Berdasarkan hasil yang di dapat dari 40 data masyarakat yang di uji maka di hasilkan keterangan *sanguin* 17, *koleris* 6, *melankolis* 4, dan *plegmatis* 13.

Kunci : Decision Tree, C.45, Klasifikasi, Kepribadian.

1. PENDAHULUAN

Kepribadian menjadi salah satu hal yang sangat berpengaruh dalam keberhasilan dan kegagalan seseorang dalam kehidupan baik dalam bidang pekerjaan, maupun dalam tindakan di lingkungan masyarakat. Salah satu

teori kepribadian adalah yang dikembangkan oleh *Hippocrates* dan *Galenus* yang membagi manusia atas empat golongan yaitu *sanguinis*, *melankolis*, *phlegmatis*, dan *koleris* (Fitria, 2013). selama ini di desa Ujung Serdang Tanjung Morawa belum adanya klasifikasi kepribadian pada masyarakat setempat, di

sebabkan kurangnya perhatian pemerintah setempat dalam menganalisis kepribadian masyarakat tersebut. Akibat dari persoalan ini banyak masyarakat yang suka mengkritik kinerja pemerintahan, tidak taat akan aturan dari pemerintah dan melawan aturan yang di buat pemerintah setempat atau bahkan sampai melakukan kekerasan sehingga menimbulkan keresahan terhadap masyarakat lain.

Pada penelitian sebelumnya yang berjudul klasifikasi karakteristik kepribadian mahasiswa di universitas Amikom Yogyakarta dengan menggunakan metode *Naive Bayes*. Tujuan dalam penelitian ini adalah untuk mengetahui tipe kepribadian mahasiswa. Hal ini bermanfaat bila suatu hari terjadi masalah terhadap mahasiswa tertentu, maka bidang kemahasiswaan sudah memahami keputusan/tindakan yang yang harus dilakukan untuk menangani mahasiswa tersebut sesuai dengan karakter kepribadiannya (Alam, Dimas Midyan et al., 2020). Keterbaruan dari penelitian ini ialah studi kasus di desa ujung serdang tanjung morawa, dengan mengembangkan proses klasifikasi ini menggunakan model data mining dengan metode *decision tree* dan *algoritma C4.5* dan membentuk pohon keputusan dari hasil klasifikasi kepribadian masyarakat.

Dari hasil penelitian dan evaluasi penelitian menunjukan bahwa *Algoritma Decision Tree C4.5* akurat diterapkan untuk penentuan kepribadian manusia yang sesuai aturan dan kompleks sehingga lebih simpel dalam memprediksi kepribadian masyarakat pada desa Ujung Serdang Tanjung Morawa.

2. METODE PENELITIAN

2.1 Jenis penelitian

Pada penelitian yang digunakan untuk penulisan laporan ini adalah data sekunder. Data sekunder adalah pengumpulan data melalui cara tidak langsung atau harus melakukan pencarian mendalam dahulu seperti melalui internet, literatur, statistik, buku, dan lain- lain (Tanujaya, 2017).

2.2 Teknik Pengumpulan Data

Teknik pengumpulan data dalam penelitian kualitatif adalah dengan melakukan wawancara dan studi studi pustaka (Prasanti, 2018) :

a. Metode Wawancara

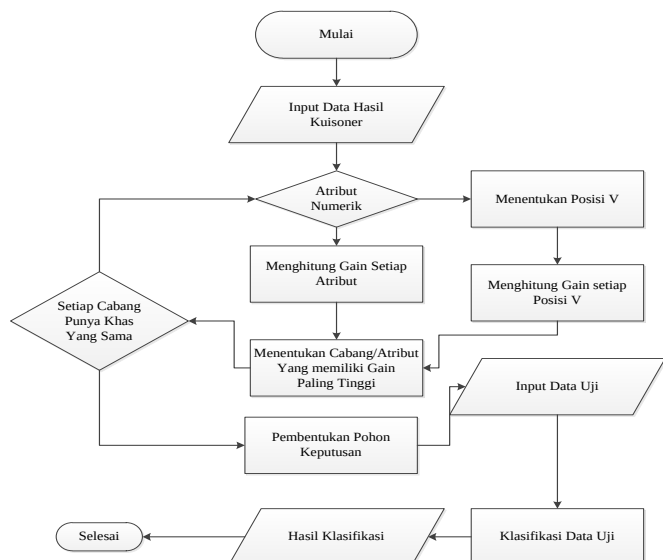
Memberikan beberapa pertanyaan dalam bentuk kuisioner dan melakukan pengolahan data dari hasil jawaban yang sudah di dapat dari jawaban hasil kuisioner.

b. Metode Studi Pustaka

Metode ini di lakukan dengan cara mencari beberapa referensi jurnal dan buku-buku yang terkait pada penelitian ini yang digunakan untuk penyelesaian masalah yang ada pada penelitian ini.

2.3 Perancangan Sistem

Perancangan sistem yang di lakukan adalah pembuatan sistem yang dapat menganalisis kepribadian manusia secara ilmu psikologi dengan mengimplemtasikan metode *decision tree C4.5*. Perancangan sistem terdiri dari perancangan arsitektur, perancangan komponen, perancangan data, serta perancangan antarmuka (Raharjo et al., 2019). untuk membentuk suatu pohon keputusan yang dapat menentukan kepribadian manusia pada desa ujung serdang kecamatan tanjung morawa dimana perancangan sistem ini meliputi *flowchart* diagram. Urutan perintah program pada suatu diagram merupakan pengertian *flowchart* (Fuadi et al., 2020). *flowchart* dari sistem yang akan dikembangkan adalah sebagai berikut :



Gambar 1. Flowcart Algoritma C4.5

2.4 Pengujian Program

Pengujian perangkat lunak merupakan sebuah proses pengujian program yang dimaksudkan untuk mencari kesalahan pada software (Debiyanti et al., 2020). Pengujian dilakukan untuk mengetahui apakah perangkat lunak telah berjalan sesuai dengan tujuan yang ingin dicapai dengan menggunakan metode *black box testing* dan memperbaiki kesalahan (*bugs*) yang muncul. Selama pengujian dilakukan. Hasil penilaian dari beberapa peserta dapat dikombinasikan. Untuk memperoleh informasi mengenai kenyamanan pemakaian sistem. Selain itu juga akan dilakukan pengujian tingkat akurasi rekomendasi dengan menggunakan metode *decision* dan *algoritma C4.5* yang dilakukan memprediksi kepribadian masyarakat.

2.5 Data Mining

Data mining merupakan sebuah inti dari proses KDD, meliputi dugaan algoritma yang mengeksplor data, membangun model dan menemukan pola yang belum diketahui.

KDD bersifat otomatis, dapat didefinisikan sebagai pengorganisasian proses untuk pengidentifikasian yang benar, berguna dan penemuan pola dari kumpulan data yang besar dan kompleks (Putra & Wadisman, 2018).

2.6 Algoritma Decision Tree

Decision Tree (Pohon Keputusan) adalah pohon dimana setiap cabangnya menunjukkan pilihan diantara sejumlah alternatif pilihan yang ada, dan setiap daunnya menunjukkan keputusan yang dipilih (Setiawati et al., 2016).

2.7 Pengertian Website

website adalah kumpulan dari halaman situs dan dokumen yang tersebar di beberapa komputer server yang berada di seluruh penjuru dunia dan terhubung menjadi satu jaringan melalui jaringan yang disebut internet.

2.8 Pengertian dari Kepribadian Manusia

Istilah kepribadian (*personality*) berasal dari kata latin "persona" yang berarti "topeng". Pada masa Yunani Kuno para aktor memakai topeng untuk menyembunyikan identitas mereka dan untuk memungkinkan mereka memerankan tokoh dalam drama. Teknik drama ini kemudian diambil alih oleh bangsa Roma dan dari merekalah kita mendapatkan istilah "personality" atau kepribadian (Fitriani, 2011).

3. HASIL DAN PEMBAHASAN

3.1 Analisis Masalah

Implementasi data mining untuk menentukan klasifikasi kepribadian masyarakat untuk mengetahui kepribadian masyarakat berdasarkan kriteria-kriteria tertentu. Data sampel telah diambil dengan menggunakan metode *decision Tree* dan *Algoritma C4.5* untuk mengetahui kepribadian masyarakat pada desa ujung serdang Tanjung Morawa.

3.2 Analisis Data

Data yang digunakan untuk klasifikasi kepribadian masyarakat pada desa ujung serdang Tanjung Morawa, Pengisian kuisioner yang dilakukan masyarakat. Berikut ini data yang diperoleh dari hasil melakukan kuisioner.

Tabel 1. Data hasil pengisian Kuisioner

No	Nama Field	Keterangan
1	Nama	Nama masyarakat yang mengisi Kuisioner
2	Jenis Kelamin	Jenis Kelamin Masyarakat
3	Usia	Keterangan usia Masyarakat
4	Pendidikan	Jenjang pendidikan masyarakat
5	Jawaban dari Kuisioner	Hasil Jawaban dari kuisioner
6	Keterangan	Keterangan mengenai klasifikasi kepribadian masyarakat

Dari data diatas yang dipilih untuk menjadi atribut untuk menentukan kepribadian masyarakat : jenis kelamin, usia, pendidikan dan hasil jawaban kuisioner yang dilakukan masyarakat, sehingga hasil yang didapat pada

pengisian kuisioner (*Sanguins*, *Koleris*, *Melankolis*, *Plegmatis*).

Tabel 2. Data Atribut

No	Nama Atribut	Penjelasan	Rentang Nilai
1	Jenis Kelamin	Jenis kelamin masyarakat	Laki-laki, Perempuan
2	Usia	Usia masyarakat	30 - 32 Tahun
3	Pendidikan	Pendidikan masyarakat	SD, SMP, SMA, Sarjana, Magister
4	Pengelompokan Jawaban Kuisioner	Menjumlah tiap-tiap jawaban yang sudah dipilih	Jawaban A “≤ / >” (5, 10, 15, 20) Jawaban B “≤ / >” (5, 10, 15, 20) Jawaban C “≤ / >” (5, 10, 15, 20) Jawaban D “≤ / >” (5, 10, 15, 20)
5	Kelas	Kelas awal hasil kuisioner	<i>Sanguins</i> , <i>Koleris</i> , <i>Melankolis</i> , <i>Plegmatis</i>

Dari 40 data kuisioner karakteristik kepribadian yang diperoleh dari hasil pengisian kuisioner oleh masyarakat Ujung serdang Tajung Morawa, data tersebut akan di uji melalui sistem dengan menerapkan metode *decision tree* dan *algoritma C4.5*, sebagai data latih dengan komposisi data : sedangkan data yang akan diujikan *Gain = Ee* digunakan pada Implementasi Metode *Decision Tree* dan *Algoritma C4.5*.

3.3 Perhitungan Data Latih

Untuk menghitung data latih digunakan metode *Decision Tree* dan *Algoritma C4.5*. data latih ialah data hasil proses *preprocessing* yang akan diubah menjadi sebuah *tree* untuk menentukan klasifikasi kepribadian masyarakat. Pada tahaan ini atribut yang akan di jadikan *root node* yaitu dengan menghitung nilai *gain* dan *entropy*. *entropy* adalah ukuran dari teori

informasi yang dapat mengetahui karakteristik dari impurity dan homogeneity dari kumpulan data (Arifin & Fitrianah, 2018). Tapi sebelum kita menghitung nilai *gain*, harus dilakukan perhitungan *entropy*. Rumus untuk menghitung nilai *entropy* untuk semua data :

$$Entropy(S) = -\frac{w}{n} * \log_2 \left(\frac{w}{n} \right) - \frac{x}{n} * \log_2 \left(\frac{x}{n} \right) - \frac{y}{n} * \log_2 \left(\frac{y}{n} \right) - \frac{z}{n} * \log_2 \left(\frac{z}{n} \right)$$

Keterangan :

S = Atribut Entropy

w = total data *Sanguin*

x = total data *Koleris*

y = total data *Melankolis*

z = total data *Plagmatis*

n = total data keseluruhan

$$Entropy(S) = -\frac{17}{40} * \log_2 \left(\frac{17}{40} \right) - \frac{6}{40} * \log_2 \left(\frac{6}{40} \right) - \frac{4}{40} * \log_2 \left(\frac{4}{40} \right) - \frac{13}{40} * \log_2 \left(\frac{13}{40} \right)$$

$$= 0.524647733 + 0.410544839 + 0.332192809 + 0.526983722$$

$$= 1,794$$

Setelah nilai *entropy* di dapatkan maka dilakukan perhitungan *gain*. Rumus untuk menghitung nilai *gain* ialah sebagai berikut :

$$\begin{aligned} entropy(S) &= -\frac{nL}{n} * \left(-\frac{nLw}{nL} \right) * \\ &\log_2 \left(\frac{nLw}{nL} \right) - \frac{nLx}{nL} * \log_2 \left(\frac{nLx}{nL} \right) - \\ &\frac{nLy}{nL} * \log_2 \left(\frac{nLy}{nL} \right) - \frac{nLz}{nL} * \log_2 \\ &\left(\frac{nLz}{nL} \right) + \frac{nP}{n} * \left(-\frac{nPw}{nP} \right) * \log_2 \\ &\left(\frac{nPw}{nP} \right) - \frac{nPx}{nP} * \log_2 \left(\frac{nPx}{nP} \right) - \frac{nPy}{nP} \\ &\log_2 \left(\frac{nPy}{nP} \right) - \frac{nPz}{nP} * \log_2 \left(\frac{nPz}{nP} \right) \end{aligned}$$

Keterangan :

Entropy(S) = total perhitungan semua data
entropy

n = total data keseluruhan

nL = total data keseluruhan jenis
kelamin laki - laki

nP = total data keseluruhan jenis
kelamin Perempuan

nLw = total data *Sanguin* yang jenis
kelamin laki - laki

nLx = total data *Koleris* yang berjenis
kelamin laki - laki

nLy = total data *Melankolis* yang
berjenis kelamin laki - laki

nLz = total data *Plegmatis* yang berjenis
kelamin laki - laki

nPw = total data *Sanguin* yang jenis
kelamin perempuan

nPx = total data *Koleris* yang berjenis
kelamin perempuan

nPy = total data *Melankolis* yang berjenis
kelamin perempuan

nPz = total data *Plegmatis* yang berjenis
kelamin perempuan.

$$\begin{aligned} \text{Gain} &= 1,794 - \frac{17}{40} * \left(-\frac{8}{17} \right) * \log_2 \left(\frac{8}{17} \right) - \frac{1}{17} * \\ &\quad \log_2 \left(\frac{1}{17} \right) - \frac{1}{17} * \\ &\quad \log_2 \left(\frac{1}{17} \right) - \frac{6}{17} * \log_2 \\ &\quad \left(\frac{6}{17} \right) + \frac{23}{40} * \left(-\frac{8}{23} \right) \\ &\quad * \log_2 \left(\frac{8}{23} \right) - \frac{5}{23} * \\ &\quad \log_2 \left(\frac{5}{23} \right) - \frac{3}{23} * \\ &\quad \log_2 \left(\frac{3}{23} \right) - \frac{24}{23} * \log_2 \\ &\quad \left(\frac{24}{23} \right) \\ &= 1,794 - (0,731 + 0.894) \\ &= \mathbf{0,169} \end{aligned}$$

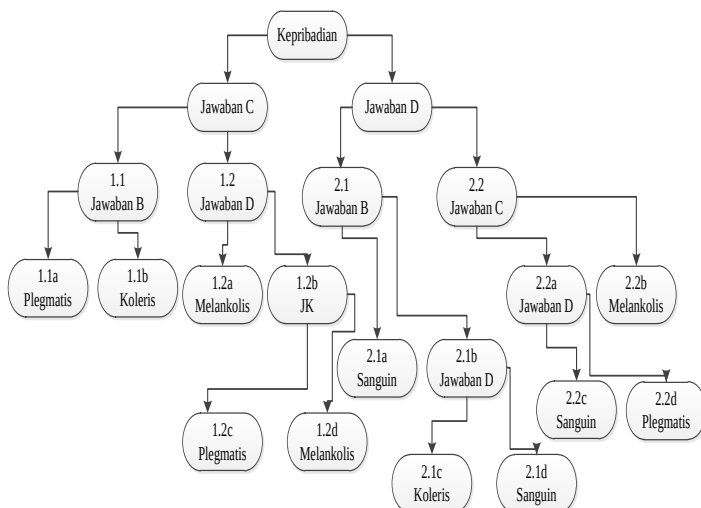
Setelah di dapat nilai *entropy* dan *gain* sudah didapatkan maka didapatkan hasil perhitungan *numeric* dari jawaban pertanyaan pada kuisioner yaitu jumlah jawaban A, jumlah jawaban B, jumlah jawaban C dan jumlah jawaban D. hasil perhitungan atribut *numeric* pada tabel dibawah ini.

Table 3. Hasil Perhitungan Atribut

No	Nilai Atribut	Jumlah Data	Sanguin	Koleris	Melankolis	Plegmatis	Entropy	Gain
1	jenis kelamin = L	17	9	1	1	6	1.497	0.057
2	jenis kelamin = P	23	8	5	3	7	1.914	
3	usia = 32	17	3	6	3	5	1.933	
4	usia = 31	17	11	0	0	6	0.937	0.355
5	usia = 30	6	3	0	1	2	1.459	
6	jawaban a <= 5	7	5	0	0	2	0.863	
7	jawaban a > 5	33	12	6	4	11	1.875	0.096
8	jawaban a <= 10	30	14	3	3	10	1.706	
9	jawaban a > 10	10	3	3	1	3	1.895	
10	jawaban a <= 15	39	17	6	4	12	1.798	0.041
11	jawaban a > 15	1	0	0	0	1	0	
12	jawaban a <= 20	40	17	6	4	13	1.794	
13	jawaban a > 20	0	0	0	0	0	0	0
14	jawaban b <= 5	1	0	1	0	0	0	
15	jawaban b <= 5	39	17	5	4	13	1.767	
16	jawaban b <= 10	14	4	5	1	4	1.835	0.131
17	jawaban b > 10	26	13	1	3	9	1.57	

18	jawaban b <= 15	36	17	6	4	10	1.754	0.134
19	jawaban b > 15	4	0	0	1	3	0.811	
20	jawaban b <= 20	39	17	6	4	12	0.798	
21	jawaban b > 20	1	0	0	0	1	0	0.041
22	jawaban c <= 5	1	0	1	0	0	0	
23	jawaban c > 5	39	17	5	4	13	1.767	
24	jawaban c <= 10	29	11	5	3	10	1.836	0.018
25	jawaban c > 10	11	6	1	1	3	1.617	
26	jawaban c <= 15	39	16	6	4	13	1.808	
27	jawaban c > 15	1	1	0	0	0	0	0.031
28	jawaban c <= 20	40	17	6	4	13	1.794	
29	jawaban c > 20	0	0	0	0	0	0	
30	jawaban d <= 5	6	2	0	1	3	1.459	0.053
31	jawaban d > 5	34	15	6	3	10	1.791	
32	jawaban d <= 10	25	10	2	3	10	1.716	
33	jawaban d > 10	15	7	4	1	3	1.746	0.067
34	jawaban d <= 15	36	16	4	4	12	1.753	
35	jawaban d > 15	4	1	2	0	1	1.5	
36	jawaban d <= 20	40	17	6	4	13	1.794	0.066
37	jawaban d > 20	0	0	0	0	0	0	
38	jawaban a <= 20	40	6	4	13	17	1.794	
39	jawaban a > 20	0	0	0	0	0	0	0
40	jawaban b <= 20	40	17	6	4	13	1.794	

Dari hasil perhitungan data diatas di dapatlah pohon keputusan yang digunakan untuk klasifikasi kepribadian pada masyarakat berdasarkan hasil perhitungan atribut *numeric*.



Gambar 2. Pembentukan pohon keputusan

Setelah semua atribut sudah di hitung dan di dapatkan nilai *entropy* dan *gain* didapatkan aturan *rule* dalam bentuk *IF* dan *THEN* sebagai berikut :

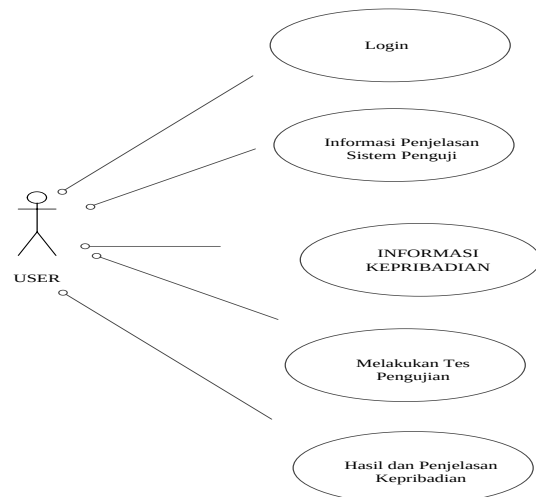
1. *IF* (usia='30') *AND* (jawaban_c<=10) *AND* (jenis_kelamin='L') *AND* (jawaban_a<=10) *THEN* Label = *Sanguin*.
2. *IF* (usia='30') *AND* (jawaban_c<=10) *AND* (jenis_kelamin='L') *AND* (jawaban_a>10) *THEN* Label = *Plegmatis*.
3. *IF* (usia='30') *AND* (jawaban_c<=10) *AND* (jenis_kelamin='P') *THEN* Label = *Plegmatis*.
4. *IF* (usia='30') *AND* (jawaban_c>10) *THEN* Label = *Sanguin*.
5. *IF* (usia='32' *OR* usia='31') *AND* (usia='32') *AND* (jawaban_d<=10) *AND* (jawaban_a<=10) *AND* (jawaban_b<=10) *AND*

- (jenis_kelamin='L') THEN Label = Plegmatis.
6. IF (usia='32' OR usia='31') AND (usia='32') AND (jawaban_d<=10) AND (jawaban_a<=10) AND (jawaban_b<=10) AND (jenis_kelamin='P') THEN Label = Sanguin.
7. IF (usia='32' OR usia='31') AND (usia='32') AND (jawaban_d<=10) AND (jawaban_a<=10) AND (jawaban_b>10) AND (jawaban_b<=15) AND (jenis_kelamin='L') THEN Label = Plegmatis.
8. IF (usia='32' OR usia='31') AND (usia='32') AND (jawaban_d<=10) AND (jawaban_a<=10) AND (jawaban_b>10) AND (jawaban_b<=15) AND (jenis_kelamin='P') THEN Label = Melankolis.
9. IF (usia='32' OR usia='31') AND (usia='32') AND (jawaban_d<=10) AND (jawaban_a<=10) AND (jawaban_b>10) AND (jawaban_b>15) THEN Label = Plegmatis.
10. IF (usia='32' OR usia='31') AND (usia='32') AND (jawaban_d<=10) AND (jawaban_a>10) AND (jawaban_b<=10) THEN Label = Koleris.
11. IF (usia='32' OR usia='31') AND (usia='32') AND (jawaban_d<=10) AND (jawaban_a>10) AND (jawaban_b>10) AND (jawaban_b<=15) THEN Label = Sanguin.
12. IF (usia='32' OR usia='31') AND (usia='32') AND (jawaban_d<=10) AND (jawaban_a>10) AND (jawaban_b>10) AND (jawaban_b>15) THEN Label = Melankolis.

Setelah sudah dilakukan perhitungan dengan aturan rule dengan IF dan THEN hasil proses prediksi klasifikasi kepribadian masyarakat desa ujung serdang tanjung morawa maka di dapat hasil dari 40 data masyarakat yang di uji menghasilkan keterangan sanguin 17, koleris 6, melankolis 4, dan plegmatis 13.

3.4 Perancangan Sistem

Perancangan sistem ialah impelmentasi dari ide ide yang ada di pikirkan untuk mengatasi permasalahan dan dituangkan ke dalam metode *decision* dan *algoritma C45* klasifikasi kepribadian masyarakat di desa ujung serdang tanjung morawa, seperti *Use Case Diagram* di bawah in:



Gambar 3. Use Case Diagram Aplikasi Klasifikasi Kepribadian

3.5 Implementasi Sistem

Implementasi sistem ialah berbasis website dari rancangan dari proses analisis dan perancangan sistem. Tampilan antar muka pada sistem klasifikasi kepribadian masyarakat antara lain : halaman login, halaman kuisoner, halaman data mining, halaman hasil uji pohon keputusan.

3.5.1 Halaman Login

Halaman login merupakan halaman yang digunakan untuk masuk ke dalam sistem, dengan memasukkan username dan password, seperti gambar di berikut :

Gambar 4. Halaman Login

3.5.2 Halaman Utama (Home)

Halaman utama (*home*) tampilan awal sistem implementasi metode decision algoritma C45 klasifikasi kepribadian masyarakat.



Gambar 5. Halaman Home

3.5.3 Halaman Pengisian Kuisioner

Halaman pengisian kuisioner ialah tampilan di mana peserta bias login sendiri ke dalam sistem dan menjawab pertanyaan yang ada untuk menentukan kepribadiannya masing-masing.



Gambar 6. Halaman Pengisian Kuisioner

3.5.4 Halaman Proses Mining

Halaman ini ialah halaman pengolahan data uji ke proses mining untuk mendapatkan nilai entropy dan gain pada tiap atribut.

Gambar 7. Halaman Proses Mining

3.5.5 Halaman Uji Pohon Keputusan

Halaman Uji pohon keputusan ialah proses pengujian data setelah dilakukan proses mining dan mendapatkan nilai entropy dan gain yang dapat digunakan untuk menentukan aturan *rule* dengan *IF* dan *THEN*.

Gambar 8. Halaman Uji Pohon Keputusan

3.5.6 Halaman Hasil Klasifikasi

Pada halaman ini kita dapat melihat hasil prediksi dari klasifikasi kepribadian masyarakat, seperti gambar di bawah berikut :

Uji Pohon Keputusan

No	Nama	LP	Usia	Pendidikan	Jenis Kelamin	Jenis Kelamin	Jenis Kelamin	Jenis Kelamin	Referensi	Referensi	Referensi	Referensi
1	Ruben	P	22	SD	11	11	11	11	Sanguin	Sanguin	11	Benar
2	Rubi	P	21	SD	11	11	11	11	Sanguin	Pegmat	11	Salah
3	Andi	P	22	SD	4	11	11	11	Pegmat	Pegmat	11	Benar
4	Ruben	P	22	SD	11	11	11	11	Sanguin	Sanguin	11	Benar
5	Unggan	P	22	SD	11	11	11	11	Sanguin	Pegmat	11	Salah
6	Unggan	P	22	SD	11	11	11	11	Sanguin	Pegmat	11	Salah
7	Michael	P	21	SD	11	11	11	11	Sanguin	Pegmat	11	Salah

Gambar 9. Halaman Hasil Klasifikasi

4. SIMPULAN

Implementasi data mining dibentuk untuk menganalisis kepribadian masyarakat secara efektif dan akurat, dan bahkan dapat melakukan pengolahan data dalam jumlah besar. Dari hasil analisis data masyarakat di desa ujung serdang, maka ada 4 atribut yang ditetapkan sebagai prediktor yaitu jenis kelamin, usia, pendidikan dan kuisoner yang berupa 40 pertanyaan. Hasil dari klasifikasi yang didapat dari analisis ialah sebagai berikut sanguin 17, koleris 6, melankolis 4, dan plegmatik 13 dan pada Implementasi sistem menunjukkan bahwa nilai entropy tertinggi ialah 1,794 dan nilai gain tertinggi 0,169.

5. DAFTAR PUSTAKA

Alam, Dimas Midyan, J. I., No, I. I., & Pratama, M. W. (2020). *Jurnal Jendela Inovasi Daerah DENGAN MENGGUNAKAN METODE NAIVE BAYES. III*, 17–28.

Arifin, M. F., & Fitriana, D. (2018). Penerapan Algoritma Klasifikasi C4.5 Dalam Rekomendasi Penerimaan Mitra Penjualan Studi Kasus: PT Atria Artha Persada. *InComTech*, 8(2), 87–102. <https://doi.org/10.22441/incomtech.v8i1.12198>

Debiyanti, D., Sutrisna, S., Budrio, B., Kamal, A. K., & Yulianti, Y. (2020). Pengujian Black Box pada Perangkat Lunak Sistem Penilaian Mahasiswa Menggunakan Teknik Boundary Value Analysis. *Jurnal*

Informatika Universitas Pamulang, 5(2), 162.

<https://doi.org/10.32493/informatika.v5i2.5446>

Fitria. (2013). 濟無No Title No Title. *Journal of Chemical Information and Modeling*, 53(9), 1689–1699.

Fitriani, A. (2011). Pendekatan “empat p” dalam kegiatan pendidikan dan pembelajaran. *Pendekatan “Empat P” Dalam Kegiatan Pendidikan Dan Pembelajaran*, Vol. 9(No. 1), hal. 58.

Fuadi, S., Candra, O., Padang, U. N., Prof, J., & Air, H. (2020). Prototype Alat Penyiram Tanaman Otomatis dengan Sensor Kelembaban dan Suhu Berbasis Arduino. *Jurnal Teknik Elektro Indonesia*, 1(1), 21–25.

Prasanti, D. (2018). Penggunaan Media Komunikasi Bagi Remaja Perempuan Dalam Pencarian Informasi Kesehatan. *LONTAR: Jurnal Ilmu Komunikasi*, 6(1), 13–21. <https://doi.org/10.30656/lontar.v6i1.645>

Putra, R. R., & Wadisman, C. (2018). Implementasi Data Mining Pemilihan Pelanggan Potensial Menggunakan Algoritma K Means. *INTECOMS: Journal of Information Technology and Computer Science*, 1(1), 72–77. <https://doi.org/10.31539/intecom.v1i1.141>

Raharjo, D. I. N. Rizki P., Kurniawan, T. A., & Rusdianto, D. S. (2019). Pengembangan Aplikasi Penentuan Prioritas Kebutuhan Fungsional Perangkat Lunak Berdasarkan Kebutuhan Non-Fungsional. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 3(3), 2862–2866.

Setiawati, D., Taufik, I., Jumadi, J., & Zulfikar, W. B. (2016). Klasifikasi Terjemahan Ayat Al-Quran Tentang Ilmu Sains Menggunakan Algoritma Decision Tree Berbasis Mobile. *Jurnal Online Informatika*, 1(1), 24. <https://doi.org/10.15575/join.v1i1.7>

Tanujaya, C. (2017). Perancangan Standart Operational Procedure Produksi Pada Perusahaan Coffein. *Jurnal Manajemen Dan Start-Up Bisnis*, 2(1), 90–95.